

Historic, archived document

Do not assume content reflects current scientific knowledge, policies, or practices.

Contract No.: FNS 53-3198-8-95 (Task Order 6)
MPR Reference No.: 8031-300

**RECOMMENDATIONS ON SOURCES OF
SURVEY DATA ON FOOD CONSUMPTION
TO MEET FNS' RESEARCH NEEDS**

August 28, 1992

Authors:

**Thomas Fraker
Sheena McConnell**

Submitted to:

**United States Department of Agriculture
Food and Nutrition Service
3101 Park Center Drive
Alexandria, VA 22302**

Attention: Pat Dinkelacker

Submitted by:

**Mathematica Policy Research, Inc.
600 Maryland Avenue, S.W.
Suite 550
Washington, D.C. 20024**

Project Director: Thomas Fraker

ACKNOWLEDGMENTS

The authors would like to thank Pat Dinkelacker, Ted Macaluso, and Margaret Andrews of the Food and Nutrition Service, Mary Hama of the Human Nutrition Information Service (HNIS), and Steve Lutz, formerly of HNIS, for their extremely helpful comments on an earlier version of the first four chapters of this report. In addition, Mary Hama provided extremely helpful information about the Nationwide Food Consumption Survey and other HNIS databases, and Steve Lutz provided a detailed description of the Thrifty Food Plan. Professor David Eastwood of The University of Tennessee made a major contribution as a co-author of a section of the report on grocery store scan data (Section III.H). Pat Doyle and Jim Ohls of Mathematica Policy Research, Inc. provided comments on an earlier draft of this report. Daisy Ewell provided excellent research assistance. We also thank Daryl Hall who edited the report and Chiquita Payne who prepared the manuscript.

CONTENTS

Chapter	Page
	ACKNOWLEDGMENTS iii
	EXECUTIVE SUMMARY xiii
I	INTRODUCTION 1
	A. TYPES OF DATA ON FOOD CONSUMPTION 1
	1. Food Expenditure 2
	2. Food Use 3
	3. Food Intake 3
	B. FNS' NEEDS FOR SURVEY DATA ON FOOD CONSUMPTION 4
	1. The Impacts of FNS' Programs on Food Consumption 5
	2. Program Design 5
	3. Program Participation 5
	4. Descriptive Studies of Food Consumption 6
	5. Methodological Studies 6
	C. THE 1987-88 NATIONWIDE FOOD CONSUMPTION SURVEY 6
	D. ALTERNATIVE SOURCES OF DATA ON FOOD CONSUMPTION 8
	E. THE STRUCTURE OF THIS REPORT 9
II	FNS' NEEDS FOR SURVEY DATA ON FOOD CONSUMPTION 11
	A. EVALUATING THE IMPACT OF FNS' PROGRAMS ON FOOD CONSUMPTION 14
	1. Estimating the Impact of the FSP on Household Food Use 15
	2. Estimating the Impact of WIC on the Diets of Women, Infants, and Children 19
	3. Evaluating the Effectiveness of Nutrition Education Programs 23
	B. PROGRAM DESIGN 24
	1. Revising the Thrifty Food Plan (TFP) 24
	2. Estimating Economies-of-Scale Adjustments in FSP Benefits 29
	3. Changing the FSP Benefit Issuance System, Criteria for Authorizing Retail Stores, and Coupon Denominations 32

CONTENTS (continued)

Chapter	Page
II (continued)	
C. MODELING PROGRAM PARTICIPATION	34
1. Modeling Participation in the NSLP and SBP	34
D. DESCRIPTIVE STUDIES OF FOOD CONSUMPTION	36
1. Describing the Food Choices of Low-Income Households	36
2. Describing the Food Intake of Low-Income Children	37
E. METHODOLOGICAL STUDIES	38
1. Estimating the Distribution of Usual Dietary Intake with Only One Day of Dietary Intake Data	38
2. Developing Methodologies for Analyzing Self-Selection in Models of Multiple Program Participation	40
3. Validating Data From the Food Stamp Cashout Demonstration Evaluations	41
III THE 1987-88 NATIONWIDE FOOD CONSUMPTION SURVEY (NFCS)	43
A. DESCRIPTION OF THE 1987-88 NFCS	43
1. Household Food Use Data	44
2. Individual Food Intake Data	45
3. Household and Individual Characteristics	45
B. POTENTIAL PROBLEMS IN THE 1987-88 NFCS	46
1. Loss of Precision Resulting from Nonresponse	47
2. Bias Resulting from Nonresponse and Incomplete Coverage	49
C. COMBINING DATA FROM THE BASIC AND LOW-INCOME SAMPLES OF THE 1987-88 NFCS	58
1. The Potential Benefits of Combining Data From the Two Samples of the 1987-88 NFCS	59
2. The General Problem of Combining Data From a Basic Sample and a Low-Income Sample	59
3. The Specific Problem of Combining Data From the Two Samples of the 1987-88 NFCS	60
D. USEFULNESS OF THE 1987-88 NFCS	61

CONTENTS (continued)

Chapter		Page
IV	ALTERNATIVE SOURCES OF DATA ON FOOD CONSUMPTION	65
A.	THE CONTINUING SURVEY OF FOOD INTAKES BY INDIVIDUALS (CSFII)	66
1.	Description of the CSFII	68
2.	Combining 1989-91 CSFII Data to Obtain More Observations	70
3.	Strengths and Weaknesses of the CSFII in Meeting FNS' Research Needs	72
B.	THE NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY (NHANES)	73
1.	Description of NHANES	73
2.	Strengths and Weaknesses of NHANES III in Meeting FNS' Research Needs	76
C.	THE CONSUMER EXPENDITURE SURVEY (CES)	77
1.	Description of the CES	77
2.	Strengths and Weaknesses of the CES in Meeting FNS' Research Needs	80
D.	THE 1977-78 NFCS AND THE 1979-80 SURVEY OF FOOD CONSUMPTION IN LOW-INCOME HOUSEHOLDS (SFC-LI)	80
1.	Differences in Survey Design	81
2.	Strengths and Weaknesses of the 1977-78 NFCS and the 1979-80 SFC-LI in Meeting FNS' Research Needs	83
E.	THE SURVEY OF INCOME AND PROGRAM PARTICIPATION (SIPP)	84
1.	Description of SIPP	84
2.	Strengths and Weaknesses of SIPP in Meeting FNS' Research Needs	85
F.	THE PANEL STUDY OF INCOME DYNAMICS (PSID)	86
1.	Description of the PSID	86
2.	Strengths and Weaknesses of the PSID in Meeting FNS' Research Needs	87

CONTENTS (continued)

Chapter	Page
IV (continued)	
G. THE DIET AND HEALTH KNOWLEDGE SURVEY (DHKS)	88
1. Description of the DHKS	88
2. Strengths and Weaknesses of the DHKS in Meeting FNS' Research Needs	89
H. GROCERY STORE SCAN DATA	90
1. General Information on Scanner Systems	91
2. Forms of Scan Data	96
3. Uses of Scan Data by Food Retailers	97
4. Proprietary Data Sets Derived from Scanner Systems	100
5. Food-Purchase Data in an EBT Environment	104
6. Strengths and Weaknesses of Scan Data in Meeting FNS' Research Needs	105
I. DATABASES FROM THE FOOD STAMP CASHOUT DEMONSTRATION EVALUATIONS	106
1. Description of the Food Stamp Cashout Demonstration Evaluations	107
2. Strengths and Weaknesses of the Food Stamp Cashout Demonstration Evaluation Databases in Meeting FNS' Research Needs	108
J. THE SCHOOL NUTRITION DIETARY ASSESSMENT (SNDA) STUDY DATABASE	109
1. Description of the SNDA Study	109
2. Strengths and Weaknesses of the SNDA Database in Meeting FNS' Research Needs	111
V FNS' NEEDS FOR SURVEY DATA: OPTIONS ON HOW TO MEET EACH NEED	113
A. EVALUATING THE IMPACT OF FNS' PROGRAMS ON FOOD CONSUMPTION	113
1. Estimating the Impact of the FSP on Household Food Use	117
2. Estimating the Impact of WIC on the Diets of Women, Infants, and Children	120
3. Evaluating the Effectiveness of Nutrition Education Programs	124

CONTENTS (continued)

Chapter	Page
V (continued)	
B. PROGRAM DESIGN	126
1. Revising the Thrifty Food Plan	126
2. Estimating Economies-of-Scale Adjustments in FSP Benefits	128
3. Changing the FSP Benefit Issuance System, Criteria for Authorizing Retail Stores, and Coupon Denominations	129
C. MODELING PROGRAM PARTICIPATION	131
1. Modeling Participation in the NSLP and SBP	133
D. DESCRIPTIVE STUDIES OF FOOD CONSUMPTION	136
1. Describing the Food Choices of Low-Income Households	136
2. Describing the Food Intake of Low-Income Children	137
E. METHODOLOGICAL STUDIES	140
1. Estimating the Distribution of Usual Dietary Intake with Only One Day of Dietary Intake Data	141
2. Developing Methodologies for Analyzing Self-Selection in Models of Multiple Program Participation	142
3. Validating Data from the Food Stamp Cashout Demonstration Evaluations	144
VI RECOMMENDATIONS ON WAYS TO MEET FNS' NEEDS FOR SURVEY DATA ON FOOD CONSUMPTION	147
A. THE 1987-88 NFCS	147
1. Background	147
2. Recommendations	148
B. ALTERNATIVE EXISTING OR PLANNED DATABASES	149
1. The 1989-91 CSFII	149
2. NHANES III	151
3. The CES	151
4. The PSID	152
5. The Cashout Demonstration Evaluation Databases	153
6. Grocery Store Scan Data	154
7. The Database for the School Nutrition Dietary Assessment Study	155

CONTENTS (continued)

Chapter	Page
VI (continued)	
C. A NEW NATIONWIDE SURVEY OF FOOD USE BY LOW-INCOME HOUSEHOLDS	155
D. RECOMMENDATIONS ON WAYS TO MEET THE DATA NEEDS OF SPECIFIC FNS RESEARCH PROJECTS	158
E. SUMMARY OF RECOMMENDATIONS	158
REFERENCES	163
APPENDIX: ABSTRACTS OF FOOD CONSUMPTION DATABASES	

TABLES

Table		Page
II.1	EXAMPLES OF FNS' RESEARCH PROJECTS AND THEIR MAJOR DATA NEEDS	12
II.2	MINIMUM SAMPLE SIZES REQUIRED TO DETECT DIFFERENCES IN FOOD USE AND NUTRIENT AVAILABILITY BETWEEN FSP HOUSEHOLDS AND NON-FSP HOUSEHOLDS	17
II.3	MINIMUM SAMPLE SIZES REQUIRED TO DETECT DIFFERENCES IN THE INTAKE OF FOOD ENERGY BETWEEN WIC PARTICIPANTS AND NONPARTICIPANTS	22
III.1	ACTUAL AND TARGET SAMPLE SIZES OF THE 1987-88 NFCS	48
III.2	NUMBER OF HOUSEHOLDS IN THE BASIC AND LOW-INCOME SAMPLES OF THE 1987-88 NFCS BY CHARACTERISTIC OF HOUSEHOLD	50
III.3	NUMBER OF INDIVIDUALS IN THE BASIC SAMPLE OF THE 1987-88 NFCS BY CHARACTERISTIC OF INDIVIDUAL	51
IV.1	CHARACTERISTICS OF DATABASES ON FOOD CONSUMPTION	67
IV.2	COMPARISON OF SAMPLE SIZES IN THE 1977-78 NFCS AND 1987-88 NFCS	82
V.1	EXAMPLES OF FNS' RESEARCH PROJECTS, PROBLEMS OF USING THE 1987-88 NFCS FOR THOSE PROJECTS, AND ALTERNATIVE SOURCES OF DATA	114
VI.1	EXAMPLES OF FNS' RESEARCH PROJECTS AND RECOMMENDATIONS ON WAYS TO MEET THEIR DATA NEEDS	160

EXECUTIVE SUMMARY

The 1987-88 Nationwide Food Consumption Survey (NFCS) collected data on household food expenditures, the use of food from the home food supply by households over a one-week period, and the intake of food from all sources by individual household members over a three-day period. Data were collected for a low-income sample (household income no greater than 130 percent of poverty) and a basic sample (no income limit). Previous versions of the NFCS (for example, the 1977-78 NFCS) have been the basis for much of FNS' past research on food consumption by low-income households.

The Food and Nutrition Service (FNS) had anticipated carrying out a broad research agenda on the basis of the 1987-88 NFCS. However, high rates of sample nonresponse to this survey have raised doubts about whether the resultant databases are capable of supporting this research. In the basic sample, only 38 percent of eligible households and 31 percent of eligible individuals participated in the survey. The low response rates have two implications for estimates based on the NFCS data. First, because the sizes of the basic and low-income samples are smaller than planned, estimates will be *less precise*. That is, the estimates will have higher variances and standard errors than would have been the case if more of the sampled households and individuals had participated in the survey. Smaller sample sizes also make it more difficult, and perhaps even unfeasible, to analyze subgroups of the population. Second, if respondents differ from nonrespondents in their food consumption, estimates based on the samples will be *biased*. A biased estimate of a variable systematically underestimates or overestimates the true value of the variable.

In light of the high rates of nonresponse to the 1987-88 NFCS, FNS contracted with Mathematica Policy Research (MPR) to review FNS' research needs for survey data on food consumption, assess the capacity of the 1987-88 NFCS to address those needs, and assess the capacity of alternative data sources to address those needs. This report presents the findings from our investigation of FNS' needs for data on food consumption and of the data sources that might be used to address those needs. The final chapter of this report makes a number of recommendations on how FNS should proceed to meet its needs for survey data on food consumption. The most significant of those recommendations are summarized in the following section, in order of importance.

SUMMARY OF RECOMMENDATIONS

1. We do not recommend using the 1987-88 NFCS on any research project that will inform policymakers or update parameters in FNS' programs.
2. We recommend conducting a new nationwide survey of food use by low-income households. Because the Continuing Survey of Food Intakes by Individuals (CSFII) provides adequate data on food intakes by individuals for most FNS applications, there should be no individual intake component to this new survey. The survey instrument should be designed with careful attention to the need to minimize respondent burden and hence to maximize response rates. The completed sample size should be approximately 5,000 households.

3. We recommend setting the income ceiling for the low-income samples of the 1993-96 CSFII at 185 percent of poverty to make it consistent with the income-eligibility criteria for the WIC program and the school nutrition programs. The income ceiling for the corresponding samples of the 1989-91 CSFII was 130 percent of poverty.
4. We recommend combining the samples of the 1989-91 CSFII across the three years of the survey and using the resultant databases on research projects that require data on food intake by individuals, especially those that require multiple days of intake data. The 1989 CSFII provides three days of food intake data on 2,925 individuals in the basic sample and 1,372 individuals in the low-income sample.
5. We recommend using the National Health and Nutrition Examination Survey III (NHANES III) when it is necessary to analyze the food intake of individuals in small segments of the population. NHANES III will provide one day of food intake data on 35,000 persons.
6. We recommend using multiple databases on certain research projects. For example, we recommend using data on individual food intake from the 1989-91 CSFII and data on household food use from the recommended survey of household food use (Recommendation 2, above) to revise the Thrifty Food Plan.
7. We recommend using data from program and demonstration evaluations in analyses other than those required by the evaluations. For example, FNS' model of participation in the National School Lunch Program is based on data from the National Evaluation of School Nutrition Programs. Opportunities for analogous applications of the data from the Food Stamp Cashout Studies and the School Nutrition Dietary Assessment Study are likely to arise in the future.
8. We recommend that FNS enhance its knowledge of grocery store checkout scanner systems and their associated databases and that it monitor ongoing developments in the collection, storage, dissemination, and analysis of such data.

I. INTRODUCTION

The Food and Nutrition Service (FNS) of the U.S. Department of Agriculture (USDA) administers more than a dozen food assistance programs that are designed to improve the diets and nutritional well-being of low-income Americans. FNS uses data on food consumption from a number of sources to assess the dietary status of the population groups served by these programs, to establish benefit levels and other program parameters, and to evaluate the effectiveness of the programs in improving diets. This report reviews FNS' needs for data on food consumption, discusses the limitations of a recent USDA survey that was intended to address many of those needs, examines alternative sources to which FNS might turn for such data, and makes recommendations about the data sets on which FNS should rely in addressing its needs for information on food consumption.

A. TYPES OF DATA ON FOOD CONSUMPTION

In this report, we use the term "food consumption" generically. That is, we use it to refer to an entire class of food-related behavior that consists of three components: expenditures on food by a household, the use of food by a household, and food eaten by an individual. We also use the term to refer to one or more of these components when the components are not specified. Each of these components has a very precise definition, given below, and each is used when we want to refer to only those individual and household food-related behaviors that are included in that definition.

Data on food consumption are typically obtained either from sources within the food distribution industry or through surveys of households or individuals. Data from industry sources are aggregated by food item or food group and provide information about the total quantity or dollar value of food that was purchased nationwide during a given period. Aggregate data are useful in tracking trends in food purchased over time. Data from surveys provide information about food consumption by households or individuals, as well as information about household and individual economic and sociodemographic characteristics. Because FNS' food assistance programs are designed to serve only

well-defined population groups, such as low-income households and pregnant and breastfeeding women who are members of low-income households, its needs for data on food consumption can often be best addressed by surveys of households and individuals.

A food consumption survey collects data on any of three measures of food consumption: expenditures on food by households, food used by households from their home food supplies, and food actually eaten ("food intake") by individuals. These measures differ with respect to whether the behavior that is being observed is the purchasing of food or the subsequent use or intake of that food, whether the measure includes all food that was consumed or a subset of that food, and whether the household or the individual is the unit of observation. We explain each of these measures of food consumption below.

1. Food Expenditure

The most straightforward approach to measuring food consumption is to ask households to recall or to keep a record of their purchases of food during a specified period of time. When a disaggregated measure of expenditure on food is required, households are asked to report the cost and, in some surveys, the quantity of each type of food that they purchased for use at home. When an aggregated measure of expenditure on food is sufficient, households are asked to recall the total amount spent on food for home use. Under this approach, regardless of whether an aggregated or disaggregated measure of expenditure on food used at home is required, households are also asked to report their total expenditure on food purchased and eaten away from home, such as restaurant meals.

Measuring food consumption by food expenditure is appropriate in many FNS applications because it is easily understood and acts as a good proxy for more sophisticated measures. A distinct advantage of this measure of food consumption is that the collection of data on food expenditures generally creates a smaller respondent burden than does the collection of data on other measures of

food consumption. Associated with this lower respondent burden is a relatively low cost of data collection and a higher rate of response.

2. Food Use

When food consumption is measured by food use, households are asked to report on the types, quantities, and prices of *all* food (not just purchased food) that they used from their home food supplies during a specific period, typically one week.¹ These data are converted into measures of the money value and nutritional value of the food used by the responding households. The latter are referred to as measures of "nutrient availability." The constructed measures of nutrient availability do not include nutrients provided by food purchased and eaten away from home, but they do include adjustments for the fact that the household was not fully dependent on the home food supply for all of its nutrients.

Measuring food consumption by food use is more appropriate than measuring it by food expenditure in most assessments of the adequacy of household diets. This is true for two principal reasons. First, food use is a more comprehensive measure of food consumption than is food expenditure because it includes home-produced food and food received as a gift or as payment-in-kind, whereas food expenditure does not. Second, food use measures food consumption as food is physically used by a household, whereas food expenditure measures food consumption as food is purchased; for this reason, food use better reflects the adequacy of a household's diet than does food expenditure.

3. Food Intake

Food intake data are collected at the individual level, unlike food expenditure data and food use data, which are collected at the household level. Respondents are asked either to recall all of the

¹In a survey of food use, data on a household's expenditure on food purchased and eaten away from home are usually obtained in the same manner as described above in Section 1.

types and quantities of food that they ate during a just-completed period, or to keep a diary of the food that they eat during a specified future period. The reference period for food intake surveys is typically one to three days. The food intake data are converted into measures of the nutritional value of the food that was eaten by the survey respondents. These are referred to as measures of "nutrient intake."

Food intake data provide an empirical basis on which to assess dietary adequacy that is superior to that provided by either food expenditure data or food use data. This is true for two reasons. First, measuring food consumption by food intake includes information on *all* food eaten by an individual, regardless of how the food was obtained (purchased or nonpurchased) or where it was eaten (at home or away from home). Second, this measure excludes food that was purchased or used, but not actually eaten, by an individual (for example, food that was fed to pets).

There are two principal reasons why food intake data may not be appropriate in certain FNS applications. First, the collection of detailed data on food intake imposes a large burden on survey respondents, thus resulting in high data collection costs and, possibly, a lower rate of response. Second, the largest of FNS' food assistance programs, the Food Stamp Program (FSP), is designed to stimulate the use of *purchased* food at home by *households*; consequently, because food intake measures *all* food eaten by *individuals*, it may not focus closely enough on the outcome of greatest interest in studies of the FSP.

B. FNS' NEEDS FOR SURVEY DATA ON FOOD CONSUMPTION

Timely and reliable survey data on food consumption are critical to many aspects of FNS' administration of federal food assistance programs. We identify in this section (and explain fully in Chapter II) five broad areas in FNS' research agenda for which data on food consumption are required.

1. The Impacts of FNS' Programs on Food Consumption

In order to monitor the success of its programs in achieving the goal of improving the diets of low-income individuals and households, FNS uses survey data on food consumption to evaluate the impacts of these programs on food expenditures and food use by households and on food intake by individuals. The evaluations may be based on simple comparisons between food consumption outcomes for program participants and eligible nonparticipants, or they may be based on estimates of program impacts on food consumption produced by econometric models.

2. Program Design

FNS uses food consumption data to design major components of its food assistance programs. This is particularly true for the FSP, in which the basis for benefit allotments is the most economical of the USDA's four model food plans--the Thrifty Food Plan (TFP). Data on both household food use and individual food intake were used to develop the TFP (HNIS, 1983). If the methodology that was used to develop the TFP is also used in the future to revise the food types and quantities that comprise the TFP, then that revision will require data on household food use and individual intake.

3. Program Participation

In order to more fully meet the needs of the target population groups for its programs, FNS must understand why some eligible households and individuals in those groups participate in these programs and others do not. Research on participation in these programs usually does not require food consumption data per se, but information on behavior associated with food consumption can be critical. Food consumption surveys often collect this critical information, whereas other types of surveys often do not. Research of this type includes studies of participation in school nutrition programs. Food consumption surveys often collect data on the frequency with which children receive USDA school breakfasts and lunches and the cost to the children of those meals. This information,

which is essential to the modeling of participation in school nutrition programs, is not available in many surveys that do not have as their central objective the gathering of data on food consumption.

4. Descriptive Studies of Food Consumption

FNS uses survey data on food consumption to describe the diets of the households and individuals who belong to the target population groups for its programs and to assess the nutritional adequacy of those diets.

5. Methodological Studies

Most of the research sponsored by FNS is intended to inform the policymaking process or to improve the administration of existing food assistance programs. FNS also sponsors methodological research that is intended to expand and improve its capacity to conduct policy-relevant research on the basis of survey data. Much of that methodological research focuses by necessity or by choice on food consumption survey databases.

C. THE 1987-88 NATIONWIDE FOOD CONSUMPTION SURVEY

Many surveys collect data on household food expenditures. A much smaller number of surveys collect data on household food use or individual food intake; these surveys also typically collect data on aggregate household food expenditures. Only one survey, the USDA's decennial Nationwide Food Consumption Survey (NFCS), collects data on all three food consumption measures: household food expenditures, household food use, and individual food intake. Because it so comprehensively measures food consumption, the NFCS has been the survey data set of first choice by FNS in a wide range of applications over the past two decades.

We noted in the Section A that the collection of data on household food use *or* on individual food intake can impose a large burden on survey respondents. Because the NFCS collects detailed data on both of these measures, it places an unusually heavy burden on respondents. This burden,

in conjunction with other factors associated with the substance and management of the survey, resulted in unusually low response rates to the 1987-88 edition of the NFCS.²

The low response rates to the 1987-88 NFCS have two principal adverse implications for estimates based on food consumption data collected by the survey. First, because the achieved sample sizes for the survey are smaller than the intended sample sizes, estimates will be less precise than planned, and it may not be possible to produce estimates for some population subgroups of interest. Second, if nonresponse is correlated with food consumption, then estimates of food consumption behavior by households and individuals nationwide that are based on the collected data will be biased.

In response to the low rates of participation in the 1987-88 NFCS by eligible households and individuals, the survey's sponsoring agency, the USDA's Human Nutrition Information Service (HNIS), requested that the Life Sciences Research Office (LSRO) of the Federation of American Societies for Experimental Biology review the survey and make recommendations regarding the use of the data. LSRO convened an expert panel consisting of three statisticians with expertise in sample design and survey nonresponse to conduct the review and make recommendations.³ The panel found the potential problems from nonresponse to the 1987-88 NFCS to be sufficiently serious that it "does not recommend use of the data" (LSRO, 1991, page 13).

This report draws upon both the findings of the LSRO expert panel and the findings from other evaluations of the 1987-88 NFCS that have been conducted by government agencies and academics. Chapter III of this report describes the 1987-88 NFCS, identifies its potential weaknesses and the

²The factors that are believed to have contributed to the low rates of response to the 1987-88 NFCS are discussed in Life Sciences Research Office (1991) and General Accounting Office (1991).

³The three members of the expert panel were: Richard Bolstein, Ph.D., Professor, School of Information Technology and Engineering, George Mason University; Charles Cowan, Ph.D., Chief Statistician, Opinion Research Corporation; and Ronald N. Forthofer, Ph.D., Professor, Department of Biometry, School of Public Health, University of Texas (LSRO, 1991, page 19).

work that has been done to remedy those weaknesses, and assesses its usefulness to FNS in carrying out its food consumption research agenda.

D. ALTERNATIVE SOURCES OF DATA ON FOOD CONSUMPTION

Despite the LSRO expert panel's pessimistic assessment of the utility of the 1987-88 NFCS in studies of food consumption, there are likely to be some applications in which the use of those data is appropriate. However, the limitations of the 1987-88 NFCS mean that would-be users of those data must carefully consider whether other existing or potential data sets would provide a better empirical basis for their studies.

There are five categories of alternatives to the 1987-88 NFCS. The first includes earlier editions of the NFCS, such as the 1977-78 NFCS. These surveys achieved substantially higher response rates than did the 1987-88 NFCS,⁴ but the data they collected are now more than ten years old and, consequently, are not believed to accurately represent current food consumption behavior. The second alternative is the USDA's Continuing Survey of Food Intakes by Individuals, an annual survey that collects detailed data on individual food intake and aggregate data on household food expenditure from relatively small numbers of respondents. The third category includes recent or ongoing nationwide surveys, the principal objective of which is not the collection of food consumption data, but which nevertheless do collect data on one or two food consumption measures. The National Health and Nutrition Examination Survey is one such alternative. The fourth category is data on retail food sales from sources within the food distribution industry. Several commercially available data sets from this source provide detailed data on food purchases at the household level. The fifth category is specialized surveys of individual food intake or household food use that are

⁴Sixty-one percent of sampled households participated in the 1977-78 NFCS and 57 percent of sampled individuals provided at least one day of data on food intake. The corresponding rates of participation in the 1987-88 NFCS are 38 percent for households and 31 percent for individuals. Both sets of participation rates are for the basic samples of the two surveys.

based on very restricted samples, such as the survey of food intake by school children that is being conducted under the current FNS-sponsored School Nutrition Dietary Assessment Study.

This report identifies one or more data sets in each of these categories, describes their measures of food consumption, and assesses their potential to fill the gaps created by the deficiencies in the 1987-88 NFCS.

E. THE STRUCTURE OF THIS REPORT

Chapter II of this report discusses FNS' needs for survey data on food consumption. It describes the five general areas of research for which FNS needs food consumption data and the specific types of data required within each area.

Chapter III describes the 1987-88 Nationwide Food Consumption Survey. It discusses existing reviews of the survey, the potential problems of using the data collected by the survey, and the usefulness of the survey in meeting FNS' research needs.

Chapter IV describes data sources that are, for some purposes, potential substitutes for the 1987-88 NFCS. Included in each description is an assessment of the strengths and weaknesses of each database in meeting FNS' research needs. The main features of the 1987-88 NFCS and ten alternative databases are listed in the appendix to this report.

Chapter V discusses the problems in using the 1987-88 NFCS for specific projects in the five research areas described in Chapter II and identifies other databases that may be more appropriate.

Chapter VI presents our recommendations on the use of the 1987-88 NFCS, the use of other databases, and which databases FNS should use to meet its needs for food consumption data in each of the five general research areas.

II. FNS' NEEDS FOR SURVEY DATA ON FOOD CONSUMPTION

In designing and evaluating its food assistance programs, FNS has many needs for survey data on food consumption. The purpose of this chapter is to describe these needs. In addition to discussing FNS' general needs for food consumption data, we describe the data needs of specific research projects that, when first proposed, were to have used the 1987-88 NFCS.

We have identified five research areas within FNS' overall research agenda that require survey data on food consumption:

1. Evaluating the impact of USDA food assistance programs on food consumption
2. Designing USDA food assistance programs
3. Modeling participation in USDA food assistance programs
4. Describing the food consumption of the low-income population
5. Developing new research methodologies

Strictly speaking, the modeling of participation in food assistance programs does not require food consumption data, but it does require a database with two features: a large sample of program-eligible households or individuals and detailed information on program participation and eligibility. Hence, we include the modeling of program participation in our discussion because the NFCS is one of the few databases with both of these features. We give examples of specific research projects in each of these five general areas, and for each project, we identify its particular data requirements. Table II.1 lists each of these projects and its major data needs.

FNS' specific data needs vary by research project. However, nearly all of the research projects discussed in this chapter require a database with some combination of the following characteristics:

1. It should contain some measure of food consumption. Depending on the specific application, this may be household food expenditure, household food use, or individual food intake.

TABLE II.1
EXAMPLES OF FNS' RESEARCH PROJECTS AND THEIR MAJOR DATA NEEDS

Research Project	Major Data Needs
A. Evaluating the Impact of FNS Programs on Food Consumption	
1. Estimating the Impact of the FSP on Household Use	<ul style="list-style-type: none"> • A sample of households with income not exceeding 130 percent of poverty¹ • Nutrient availability from food used at home • Proportion of meals eaten away from home • FSP participation
2. Estimating the Impact of WIC on the Diets of Women, Infants, and Children	<ul style="list-style-type: none"> • A sample of WIC-eligible individuals with household income not exceeding 185 percent of poverty • Individual food and nutrient intake • WIC participation • Whether women are pregnant, breastfeeding, or postpartum
3. Evaluating the Effectiveness of Nutrition Education Programs	<ul style="list-style-type: none"> • Longitudinal database or specialized cross-sectional database with information on participants in nutrition education programs • Knowledge and attitudes about diet • A measure of household or individual food consumption
B. Program Design	
1. Revising the Thrifty Food Plan	<ul style="list-style-type: none"> • A sample of households with income not exceeding 130 percent of poverty • Type, quantity, and money value of food used by households and household nutrient availability • Individual food intake
2. Estimating Economies-of-Scale Adjustments in FSP Benefits	<ul style="list-style-type: none"> • A sample of households with income not exceeding 130 percent of poverty • Money value and quantity of food used by households
3. Changing (1) the FSP Benefit Issuance System, (2) Criteria for Authorizing Retail Stores, and (3) Coupon Denominations	<ul style="list-style-type: none"> • A sample of food stamp recipients • Detailed food purchase data
C. Modeling Program Participation	
1. Modeling Participation in the NSLP and SBP	<ul style="list-style-type: none"> • A sample of school-aged children whose household income does not exceed 185 percent of poverty • Whether child attends a school that participates in the NSLP and SBP • Meal-price status of child • Frequency of eating school meals • Full and reduced price of meals

TABLE II.1 (continued)

Research Project	Major Data Needs
D. Descriptive Studies of Food Consumption	
1. Describing the Food Choices of Low-Income Households	<ul style="list-style-type: none"> • A sample of low-income households • Household food use and household nutrient availability • Participation in food assistance programs
2. Describing the Food Intake by Low-Income Children	<ul style="list-style-type: none"> • A sample of low-income children • Individual food and nutrient intake • Participation in food assistance programs
E. Methodological Studies	
1. Estimating the Distribution of Usual Dietary Intake with Only One Day of Dietary Intake	<ul style="list-style-type: none"> • A sample of households from the population of interest • Multiple days of individual nutrient intake data
2. Developing Methodologies for Analyzing Self-Selection in Models of Multiple Program Participation	<ul style="list-style-type: none"> • A sample of households eligible for food assistance programs • A measure of household or individual food consumption • Participation in food assistance programs
3. Validating Data from the Food Stamp Cashout Demonstration Evaluations	<ul style="list-style-type: none"> • A nationally representative sample of food stamp recipients • Type, quantity, and money value of food used by households and household nutrient availability

¹This income limit, as well as those given elsewhere in this table are rough approximations to the income eligibility criteria for various FNS programs.

2. It should contain information about eligibility for and participation in USDA food assistance programs. The largest of these programs are the Food Stamp Program (FSP), the Special Supplemental Food Program for Women, Infants, and Children (WIC), the National School Lunch Program (NSLP), and the School Breakfast Program (SBP).¹
3. It should be large enough to yield precise estimates and to allow for some subgroup analyses, especially of low-income households--the subgroup of most interest to FNS.
4. It should contain information on sociodemographic and economic characteristics of the household.
5. It should be fairly current. This is important because USDA food assistance programs and the diets of the populations that they serve are changing rapidly.

A. EVALUATING THE IMPACT OF FNS' PROGRAMS ON FOOD CONSUMPTION

The main objective of FNS' food assistance programs is to help low-income persons obtain nutritionally adequate diets. Hence, evaluating the impact of these programs on food consumption is one of FNS' most important areas of research. The types of data needed for such an evaluation vary by program. However, at a minimum, the database used in the evaluation must have the following four characteristics:

1. It must contain a large sample of households or individuals (depending on the program) who are eligible for the program. The required sample size depends on the program, the measure of food consumption, the need for subgroup analyses, and the minimum impact of the program on food consumption that FNS wishes to detect.
2. It must have information on whether the households or individuals in the sample participate in the program and enough information about the households or individuals to estimate whether they are eligible for the program.
3. It must contain a relevant measure of food consumption--household food expenditure, household food use, or individual food intake. While it would be interesting to examine the impact of a program using all three measures of food consumption, the most relevant measure of food consumption for the FSP is household food use; for the WIC and school nutrition programs, the most relevant measure is individual food intake.

¹The Nutrition Assistance Program is larger than the SBP but serves low-income households only in Puerto Rico.

4. If the evaluation is to examine the impact of the food assistance program on the proportion of the population that is nutritionally at risk (for example, by having intake of a nutrient that is less than the recommended dietary allowance), the database must contain more than one day of intake data from each individual. Multiple days of intake data are required to obtain unbiased estimates of the proportion at nutritional risk.

In this section, we discuss examples of research projects aimed at evaluating the impact of three FNS programs: (1) the impact of the FSP on food consumption, (2) the impact of the WIC program on the diets of its participants, and (3) the impact of nutrition education programs on the diets of the low-income population. For each program, we discuss the specific data needs of the research project.

1. Estimating the Impact of the FSP on Household Food Use

The main objective of the FSP is to enhance the ability of low-income households to purchase food and obtain a nutritionally adequate diet. Hence, one of FNS' more important research objectives is to measure the magnitude of the impact of the FSP on food consumption and the quality of the household's diet. With only minor exceptions, food stamps can legitimately be used only to purchase food to be consumed at home; hence, the measures of food consumption most directly affected by the FSP are the money value of food purchased by households for use at home and the availability of nutrients from food used at home. However, policymakers may also be interested in examining the impact of the FSP on individual nutrient intake.

A database used to estimate the impact of participation in the FSP on the money value of food used at home or nutrient availability should have four characteristics:

1. It must be based on a large sample of households whose gross incomes do not exceed 130 percent of poverty and, thus, satisfy the FSP's gross income eligibility requirement. If the results of the study are to be generalized to the entire U.S., the sample should be nationally representative of low-income households.

2. It must contain information on the availability of nutrients from the food used by the household at home. If the database is also to be used to measure the impact of the FSP on food expenditure and individual nutrient intake, it should contain information on the household's expenditures on food for home consumption, or on the money value of food used at home, and on food intake by the individual members of the household.
3. It must contain information on participation in the FSP, including the value of food stamp benefits.
4. It must contain information on household characteristics, including income; location; and the sex, age, and race of the household head.

Table II.2 provides estimates of the minimum sample sizes required to detect 5, 7.5, and 10 percent differences between FSP-participating households and non-FSP-participating households in (1) the money value of purchased food used at home per adult male equivalent (AME) in the household² and (2) a measure of nutrient availability (the availability of food energy, expressed as a percent of the RDA). We base our estimates of the minimum sample sizes required to detect differences in the money value of purchased food used at home on values of the means and standard deviations calculated from (1) the data collected in the Alabama Cashout Demonstration and (2) the 1977-78 NFCS; our estimates of the minimum sample sizes required to detect differences in food energy as a percentage of the RDA are based on the mean and standard deviation in the Alabama Cashout database. Our estimates are based on a 95 percent confidence, an 80 percent power,³ and a sample that contains an equal number of FSP and non-FSP households. The estimates are also based on the assumptions that the sample is a simple random sample and that a difference-in-means test is used to test for an impact of the FSP. The

²The number of adult male equivalents in a household is a measure of household size that adjusts for the differing nutritional requirements of household members in different age-sex groups.

³With a 95 percent confidence, if the FSP truly has no impact on food consumption, there is a 95 percent probability that we would correctly conclude from our sample that the program had no impact on food consumption. With an 80 percent power, if the FSP does have an impact on food consumption of an assumed size, there is an 80 percent probability that we would correctly conclude from our sample that the program had an impact on food consumption.

TABLE II.2
MINIMUM SAMPLE SIZES REQUIRED TO DETECT DIFFERENCES
IN FOOD USE AND NUTRIENT AVAILABILITY BETWEEN FSP
HOUSEHOLDS AND NON-FSP HOUSEHOLDS

Food Consumption Measure	Minimum Detectable Difference	Minimum Sample Size: Number of Households	
		One-Tailed Test	Two-Tailed Test
Money Value of Purchased Food Used at Home per AME (Alabama, mean = \$32, std. dev. = \$17) ^a	5.0 %	2,835	3,597
	7.5 %	1,260	1,599
	10.0 %	709	899
Money Value of Purchased Food Used at Home per AME (1977-78 NFCS, mean = \$22, std. dev. = \$9) ^b	5.0 %	1,589	2,016
	7.5 %	706	896
	10.0 %	397	504
Food Energy as a Percentage of the RDA (mean = 162%, std. dev. = 79%) ^c	5.0 %	2,359	2,994
	7.5 %	1,048	1,331
	10.0 %	590	748

NOTES: Figures are based on a 95 percent confidence, an 80 percent power, and an equal number of FSP households and non-FSP households. The minimum sample size is the combined number of FSP and non-FSP households.

^aMean and standard deviation of the money value of purchased food used at home per AME from a sample of food stamp recipients participating in the Alabama Cashout Demonstration (Fraker et al. 1992a, IV.4, and Fraker et al. 1992b, I.4). In this example, a 10 percent difference in the outcome measure would be \$3.20.

^bMean and standard deviation of the money value of purchased food used at home per AME from the 1977-78 NFCS. In this example, a 10 percent difference in the outcome measure would be \$2.20.

^cMean and standard deviation of average food energy from food used at home from a sample of food stamp recipients participating in the Alabama Cashout Demonstration (Fraker et al. 1992b, B.5). In this example, a 10 percent difference in the outcome measure would be 16.2 percentage points.

samples would need to be larger if they were not simple random samples, and could be smaller if regression analysis were used to test for the impact of the FSP.

The minimum required sample sizes are smaller if the size of the difference that needs to be detected is larger and if a one-tailed rather than a two-tailed test is used.⁴ The minimum required sample sizes are also smaller if the estimates are based on the mean and standard deviation of the money value of food used at home from the 1977-78 NFCS rather than the Alabama Cashout database.⁵ Using a one-tailed test, a sample of between 1,500 and 2,800 households would be required to detect a 5 percent difference between FSP households and non-FSP households in the money value of food used. A sample of about 2,400 would be sufficient to detect a 5 percent difference in food energy as a percentage of the RDA using a one-tailed test. Samples only one-fourth as large as these would be required to detect 10 percent differences in the food consumption measures (using a one-tailed test).

The most recent national estimates of the impact of the FSP on nutrient availability were made by Devaney, Haines, and Moffitt (1989) using the 1979-80 SFC-LI. Before the release of the 1987-88 NFCS response rates, FNS proposed to refine and reestimate the Devaney-Haines-Moffitt model using data from the low-income sample of the 1987-88 NFCS.

In its most basic form, the Devaney-Haines-Moffitt model consists of a simple ordinary least squares regression model.⁶ The dependent variable in the model is the availability of nutrients from food used at home by the household. This variable is scaled into "equivalent nutrient units" to take

⁴A one-tailed test is used if the analyst assumes that participation in the FSP does not have a negative impact on food consumption. A two-tailed test is appropriate if the analyst has no prior regarding the sign of the FSP's impact on food consumption.

⁵This is because the variability (measured by the ratio of the standard deviation to the mean) in the money value of purchased food used at home is smaller in the 1977-78 NFCS than in the Alabama Cashout database.

⁶Devaney, Haines, and Moffitt also estimated models that took into account that FSP participants and nonparticipants may differ in unobserved ways that are related to the probability of a household participating in the FSP. The estimates of these models were similar to the estimates made using ordinary least squares.

into account the variation in the age-sex composition of the household and the proportion of meals eaten away from home.⁷ The independent variables include the value of food stamp benefits received; household income; region of the country in which the household is located; and the sex, age, and race of the household head. The coefficient on the food stamp benefit variable provides an estimate of the marginal impact of food stamp benefits on nutrient availability.

The Devaney-Haines-Moffitt model should be estimated on a database obtained from a sample of low-income households, that provides information on household food use, FSP participation, and household characteristics. In addition, the database must contain information on the proportion of meals eaten by members of the household away from home in order to calculate the number of equivalent nutrition units in each household.

2. Estimating the Impact of WIC on the Diets of Women, Infants, and Children

The principal objective of the WIC program is to improve the diets of the individuals in low-income households who participate in the program--infants; preschool children; and pregnant, breastfeeding, and postpartum women who are at nutritional risk. To achieve this objective, WIC provides supplemental foods such as infant formula, milk or cheese, eggs, cereals, juices, and legumes either as a package or in the form of vouchers that can be used to purchase specific foods. Whether WIC meets its objective is important to policymakers deciding on the funding and design of the program. Therefore, to determine if the objective is met, FNS proposed estimating the impact of WIC on the diets of WIC-eligible women, infants, and children. As first proposed, this study was to be conducted using the 1987-88 NFCS.

The proposed analysis of the dietary impact of WIC would involve comparing the food consumption of WIC participants with the food consumption of WIC-eligible nonparticipants. The

⁷The number of equivalent nutrition units in a household is a measure of household size that adjusts for (1) the differing nutritional requirements of household members in different age-sex groups and (2) the proportion of all meals that household members eat away from home. Thus, household size in equivalent nutrition units is a measure of the number of "equivalent persons" who are fully dependent on the household food supply for food and nutrients.

analysis should control as much as possible for differences, other than WIC participation, between the two groups. To assess the differential impact of the program on the various demographic groups that it serves, the study should separately examine the differences between each group of WIC participants and each corresponding group of eligible nonparticipants--infants, children, pregnant women, breastfeeding women, and postpartum women.

Two measures of food consumption are relevant to assessing the impact of WIC on dietary intake. The first--the types and quantities of food eaten--enables us to examine whether participation in WIC increases the consumption of the types of foods provided by the WIC program. The second--nutrient intake--is a measure of the impact of WIC on the nutritional quality of the participant's diet. As participation in WIC occurs at the individual rather than household level, the outcomes should be measured at the individual level. However, it was also proposed that the study should examine the impact of WIC on household food use and household nutrient availability. These outcomes, although perhaps less important, are of interest because they will capture any impact of WIC benefits on the diets of non-WIC members of the WIC participant's family. Hence, there are a total of four outcomes of interest for this study: (1) individual food intake, (2) individual nutrient intake, (3) household food use, and (4) household nutrient availability.

Estimating the impact of the WIC program requires data on WIC eligibility, participation, and food consumption. Therefore, the database used for this project must contain a large sample of individuals eligible for the WIC program. Eligibility for the WIC program is determined by three criteria: (1) the person must be an infant, a child up to age 5, a pregnant woman, a breastfeeding woman, or a postpartum woman, (2) the person must be at nutritional risk, and (3) the person's household income must not exceed an income cutoff set by the state within the range of 100 to 185 percent of poverty. Unfortunately, few databases provide information on nutritional risk factors.

The database should also identify which persons participate in WIC. It should contain detailed information about other characteristics of participants such as age, race, ethnicity, employment, and education, as well as characteristics of the household, such as income.

Finally, the database should also contain information about individual food and nutrient intake. Ideally, it should contain more than one day of food intake data. The existence of intraindividual variation in nutrient intake--the day-to-day variation in each person's nutrient intake--causes estimates of the distribution of a person's "usual" nutrient intake based on a single day of food intake data to be biased. (A person's usual nutrient is his or her average nutrient intake that would persist over a lengthy period of time.) Thus, estimates of the proportion of the population who are nutritionally at risk that are made from estimates of this distribution will also be biased. If several days of food intake data are collected from each person, the distribution of average nutrient intake per person is still a biased estimate of the distribution of usual nutrient intake, but the bias is smaller than if the distribution had been estimated with only a single day of food intake for each person. Statistical procedures exist that adjust estimates of the distribution of average nutrient intake to remove the bias (National Research Council, 1986). However, these procedures also require more than one day of intake data from at least some persons in the sample. Thus, if FNS wishes to examine the impact of WIC on the proportion of the population who are nutritionally at risk, the database must contain more than one day of intake data.⁸

Estimated minimum sample sizes required to detect differences in the intake of food energy between WIC participants and nonparticipants are provided in Table II.3. (These sample sizes are based on the same assumptions about the confidence, power, and structure of the sample as the sample sizes reported in Table II.2.) A sample of about 1,160 WIC-eligible women is required to detect a 5 percent difference in food energy intake between WIC participants and nonparticipants using a one-tailed test. A sample of about 680 WIC-eligible children is required to detect a 5 percent

⁸This issue is discussed in more detail in Section E of this chapter.

TABLE II.3

**MINIMUM SAMPLE SIZES REQUIRED TO DETECT DIFFERENCES
IN THE INTAKE OF FOOD ENERGY BETWEEN WIC
PARTICIPANTS AND NONPARTICIPANTS**

Food Consumption Measure	Sample	Minimum Detectable Difference	Minimum Sample Size: Number of Individuals	
			One-Tailed Test	Two-Tailed Test
Intake of Food Energy as a Percentage of the RDA (mean = 82%, std. dev. = 28%) ^a	Women	5.0 %	1,162	1,474
		7.5 %	516	655
		10.0 %	290	369
Intake of Food Energy as a Percentage of the RDA (mean = 95%, std. dev. = 25%) ^a	Children	5.0 %	682	865
		7.5 %	303	385
		10.0 %	171	216

NOTES: Figures are based on a 95 percent confidence, an 80 percent power, and an equal number of WIC participants and nonparticipants. The minimum sample size is the combined number of WIC participants and nonparticipants.

^aThese means and standard deviations were taken from the 1985 CSFII (Fraker, Long, and Post 1990, B-12). In the example for women, a 10 percent difference in the outcome measure would be 8.2 percentage points. In the example for children, a 10 percent difference in the outcome measure would be 9.5 percentage points.

difference in food energy intake between children participating in WIC and nonparticipants using a one-tailed test. Samples of 290 WIC-eligible women and 171 WIC-eligible children are required to detect a 10 percent difference in food energy intake using a one-tailed test.

3. Evaluating the Effectiveness of Nutrition Education Programs

Over the past decade, there has been a growing interest in educating the U.S. population about good nutrition, food safety, and the link between diet and the risk of disease. Nutrition education is of particular interest to FNS because it may prove to be effective, as a supplement to the more traditional food assistance programs, in improving the diets of low-income Americans.⁹ Therefore, evaluating the impact of nutrition education programs on the diets of the low-income population is an important component of FNS' research agenda.

Nutrition education programs seek to improve diet by increasing knowledge about nutrition, by changing attitudes toward the importance of healthy eating, and, ultimately, by changing eating behavior. In evaluating the effectiveness of nutrition education programs, we would ideally like to examine the impact of a program on people's knowledge and attitudes about diet as well as the impact of a program on food consumption. This is because if a program is not effective, it is important to understand whether it is ineffective in changing people's knowledge and attitudes about diet or whether it is effective in changing people's knowledge and attitudes, but that these changes did not affect people's behavior. Therefore, any database used to evaluate nutrition education programs should provide information on food-related attitudes and knowledge, and on food consumption behavior.

The specific types of information required for an evaluation of nutrition education programs depend on the program to be evaluated. If the program is small in scale, affecting only a limited

⁹In 1968, the Extension Service of USDA introduced a nutrition education program, the Expanded Food and Nutrition Education Program (EFNEP), aimed at helping low-income families, especially those with young children, acquire the knowledge, skills, attitudes, and changed behavior necessary to improve their diets.

number of people in a limited number of sites, it can be evaluated by comparing, using a cross-sectional database, the knowledge, attitudes, and food consumption behavior of program participants with those of similar program nonparticipants. In addition, the survey would have to be designed to include a sufficient number of program participants.

A large-scale program that might affect all U.S. individuals, such as a nationwide media campaign, can only be evaluated using a longitudinal database. In this kind of study, the knowledge, attitudes, and food consumption behavior of people as measured before and after the introduction of the program would be compared.

B. PROGRAM DESIGN

One of the most important uses of large-scale food consumption surveys is to help design changes to the FSP and other food assistance programs. Small changes to these programs can affect millions of recipients and have a major impact on program costs. It is therefore especially important to make any program changes on the basis of reliable data. The specific data needs for program design vary by the program and the changes considered.

In this section, we discuss several examples of potential design changes to the FSP: (1) the revision of the Thrifty Food Plan, (2) a revision of the economies-of-scale adjustments in the food stamp benefit schedule, and (3) changes in the benefit issuance system, changes in the criteria used in authorizing retail stores to accept food stamps, and changes in the coupon denominations. Analyses of all of these design changes require a sample of nationally representative low-income households. The 1987-88 NFCS was expected to be the source of data for making decisions about the first two design changes.

1. Revising the Thrifty Food Plan (TFP)

The TFP is the most economical of four family food plans developed by (HNIS, 1983). It specifies the quantities of 31 groups of foods that households might use to create a nutritious, low-

cost diet. HNIS developed the TFP in 1975, and revised it in 1983. Currently, the TFP is based on the 1983 revision to the quantities of food that make up the TFP. Because Congress has mandated that the food stamp allotment schedule be based on the cost of the TFP, the design and revision of the plan have important policy implications.

a. Design of the TFP

The food stamp maximum allotment (the food stamp benefit paid to a household with no net monthly income) for a family of four is calculated from the monthly cost of the TFP for a household consisting of a man and a woman both between 20 and 50 years of age, a child between 6 and 8 years of age, and a child between 9 and 11 years of age. The maximum allotment is adjusted only once a year (in October), and the adjustment is based on the cost of the TFP in the previous June. Currently, the allotment is set at 103 percent of the cost of the TFP to compensate for the four-month lag in the inflation adjustment and for inflation during the succeeding year. The allotment for households that do not contain four persons is calculated by adjusting the four-person household allotment by factors that reflect economies of scale in household food use.

The TFP is based on the quantities of specific types of food in each of 31 food groups that met three specific requirements and that were most consistent with the actual food use patterns of low-income households. The idea was that the consumption patterns suggested by the plan should not require a radical change in the eating habits of low-income households. More specifically, the TFP was designed to meet three requirements:

1. **Nutrition Requirement.** The current TFP was designed to meet 105 percent of the 1980 Recommended Dietary Allowances (RDAs) as established by the National Academy of Sciences, Food and Nutrition Board (National Research Council, 1980). The 5 percent margin in excess of the RDAs compensates for the fact that because of spoilage and plate waste, not all food used by a household is eaten. The TFP was also designed to meet dietary guidelines, circa 1983, for four food components that did not have established RDAs: fat, cholesterol, caloric sweeteners, and sodium. These guidelines were proposed by various authoritative groups, including the Food and Nutrition Board (National Research Council, 1980) and USDA/DHHS (1980).

2. **Consumption Requirement.** The types and quantities of food in the TFP were required to meet certain minimum and maximum requirements set for practical and nutritional purposes. For example, the quantities of fruit and vegetables in the TFP were required to be no lower than the average actual consumption of fruit and vegetables by low-income households. The TFP was also required to include a wide variety of foods--specifically, it was required to include foods from each of the 31 TFP food groups.
3. **Cost Requirement.** To ensure continuity in the cost of the TFP, the cost of the revised 1983 TFP was required to be equal to the cost of the previous TFP, developed in 1975, when both were calculated on the basis of January 1981 prices.

A separate TFP was developed for each of 11 age-sex categories of persons. The TFP for a household was calculated by summing the quantities of food specified in the plan for each individual in the household.

The database used to estimate the TFP should contain a large, nationally representative sample of households eligible to participate in the FSP. *Ideally*, it should also contain the following three types of data:

1. **Quantities of Food Used by Individuals.** To ensure that the quantities and types of food specified in the TFP resemble as closely as possible the actual food use patterns of low-income individuals, there should be data on the average use of specific food types by individuals in each of the 11 age-sex categories. Unfortunately, this information is not available and, consequently, must be estimated.
2. **Nutrient Value of Food Used by Individuals.** Information about the quantities of nutrients available in the food used by individuals would be needed to ensure that the food plan meets the nutritional requirements. Unfortunately, this information is not available and, consequently, must be estimated.
3. **Prices.** Prices of food used, reflecting actual food choices within each of the 31 TFP food groups, would be needed to calculate the cost of the TFP and to ensure that it meets the cost requirement.

The current TFP was estimated using the low-income sample of the 1977-78 NFCS. This sample contains 4,600 households whose monthly gross income does not exceed 130 percent of poverty and hence are income-eligible to participate in the FSP. While the 1977-78 NFCS contains data on the quantities and nutrient value of food *used by households*, it does not contain data on the quantities

and nutrient value of food *used by individuals*. However, it does contain data on the quantity (as served) of foods *eaten by individuals* over three consecutive days. These individual food intake data together with the household food use data were used to estimate the quantity and nutrient value of food *used by individuals* in each of the 11 age-sex categories. The 1977-78 NFCS also contains information on both the quantity and value of foods purchased and used by the household, from which detailed prices could be calculated. These prices were used as a baseline for calculating the cost of the TFP. The average price of each food item reported in the 1977-78 NFCS was updated from the time of the survey to January 1981 using the CPI.

b. Revision of the TFP

Each month, HNIS updates the cost of the TFP to account for changes in the prices of foods. This is done by updating the cost of the "basket" of foods that make up the TFP, using the latest available values of the Consumer Price Index (CPI).¹⁰

Approximately every ten years, HNIS revises its family food plans by changing the *quantities and types* of foods in the basket of foods that is the basis for each plan. The last revision of the TFP occurred in 1983. The 1977-78 NFCS was used in making this revision.

A revision of the TFP may be necessary if it no longer reflects the optimal nutritious, low-cost diet. This may occur because:

1. Changes may occur in the dietary guidelines and in the RDAs for some nutrients. For example, the National Research Council (NRC, 1980) previously recommended that total fat intake should not account for more than 35 percent of total caloric intake, but it now recommends (NRC, 1989) that total fat intake should not account for more than 30 percent of total caloric intake. This recommendation also appears in the current edition of *Dietary Guidelines for Americans* (USDA/DHHS, 1990)

¹⁰ The quantities of food in the 31 TFP food groups are weighted averages of over 2,400 specific types of food, whereas the CPI is defined for only 51 groups of foods. In updating the cost of the TFP, the prices of all foods within a given CPI food group are assumed to change by the same amount as the change in the CPI for that group.

2. The nutrient composition of certain foods may change over time. For example, closer trimming of some cuts of meat has reduced external fat from 1/2 inch or more to approximately 1/4 inch. These particular cuts of meat contain less fat than they did in the late 1970s.
3. Changes in nutrient detection and measurement methods due to advances in technology may result in updated nutrient composition data for certain foods whose actual nutritional content has remained constant.
4. As tastes and food prices change, and as new foods become available, foods consumed by low-income households may change.
5. Relative food prices may have changed such that the estimated cost of the TFP no longer reflects its true cost. The monthly revisions of the cost of the TFP using the CPI account for some price changes. However, the baseline cost of the TFP was calculated from prices of thousands of types of food, while the CPI contains prices for only 51 food groups. The monthly price adjustment cannot account for relative changes in prices of food items within these 51 groups.

There is currently a concern that the TFP no longer reflects the optimal nutritious low-cost diet because of the first three reasons listed above--over the past decade, dietary guidelines and RDAs have changed, the nutrient content of food has changed, and assessments of the nutritional content of food have been revised. If the types of foods consumed by low-income individuals and relative food prices have not changed appreciably over the past decade, one option would be for HNIS to revise the 1983 TFP again on the basis of the 1977-78 NFCS, using the most recent RDAs and *Dietary Guidelines for Americans*, and current data on the nutrient composition of foods.

However, if the food choices of low-income individuals or relative food prices have changed appreciably since the 1977-78 NFCS was conducted, the TFP should be reestimated using a database that contains more recent information about food prices and the quantities of food used by low-income households. If the basic structure of the TFP were to remain unchanged (for example, retaining the existing 11 age-sex categories and 31 food groups), this database would be required to have the following three characteristics:

1. It must contain an up-to-date, large sample of nationally representative individuals from low-income households. A large sample size is needed because the TFP is estimated separately for each of the 11 age-sex categories. Thus, the database

should contain enough observations to estimate the intake of food from each of the 31 TFP food groups by individuals in each of the 11 age-sex categories. (The last revision of the TFP used a sample of 4,600 households.)

2. It must contain information on the types and quantities of food used by individuals *or* on the types and quantities of food used by households and individual food intake.
3. It must contain information on the prices of food purchased for use at home.

2. Estimating Economies-of-Scale Adjustments in FSP Benefits

Larger households enjoy certain cost advantages over smaller households in food purchasing, storage, and preparation. For example, a large household can take advantage of lower unit prices associated with bulk purchases and may use this food more efficiently, that is, with less waste. As a result, although total food expenditures are greater in larger households, the *per person* food expenditures may be smaller. This reduction in per person expenditure on food as the size of the household increases is said to result from "economies of scale."

The Food Stamp Act of 1977 mandated that food stamp benefits be adjusted for household size and that the adjustment "take into account economies of scale." Maximum food stamp allotments are adjusted for household size by multiplying the per person cost of the TFP for a household of four (consisting of a man and a woman age 20 to 50 years, a child age 6 to 8 years, and a child age 9 to 11 years) by an *economies-of-scale adjustment* and multiplying this adjusted per person cost by the number of persons in the household. The economies-of-scale adjustments currently used are -10 percent for a household of seven or more, -5 percent for a household of five or six, +5 percent for a household of three, +10 percent for a household of two, and +20 percent for a household of one.

These adjustments, estimated in 1975 by Kerr and Peterkin using data from the 1965 National Household Food Consumption Survey, may no longer adequately reflect the relationship between food expenditures and household size. Over the past 25 years, changes in tastes, in the variety of foods available, and in food packaging and marketing may have changed the relationship between food expenditures and household size. Moreover, since 1975, a number of methodological

improvements to the Kerr-Peterkin model have been suggested.¹¹ Before the publication of the LSRO expert panel review of the 1987-88 NFCS, FNS planned to use the 1987-88 NFCS to estimate new models of economies of scale and to reestimate the Kerr-Peterkin model.

a. Modeling Economies of Scale

Estimating economies of scale in household food purchases and use is complicated by two factors that are correlated with household size but are unrelated to economies of scale: (1) the age and sex composition of the household and (2) household income. The age and sex composition of the household affects at-home food expenditures because children typically eat less than adults, and women typically eat less than men. The income of the household affects food expenditures because it affects the household's incentive to economize on food purchases and reduce waste. To isolate the economies of scale from these other factors, it is necessary to analyze how expenditures on food vary with household size while controlling for the effects of both household composition and income.

Two approaches to estimating economies of scale have been suggested. The first approach--the expenditure-modeling approach--was used by Kerr and Peterkin (1975) to develop the current economies-of-scale adjustments. It involves using a regression model to estimate the impact of household size on per person food expenditures. In addition to a variable that measures household size, the model includes as explanatory variables measures of the household's age-sex composition and income. To control for any differences in the quality of food chosen by households of different sizes, a measure of the nutritional quality of the food purchased, such as the number of nutrients for which food used by the household provided at least the RDA for the household members, is also included as an explanatory variable in the model. Because the amount of food eaten away from home both affects expenditures on food for use at home and is correlated with household size, the model also includes a measure of the number of meals eaten away from home. The regression coefficient on the household size variable provides an estimate of the economies-of-scale adjustment.

¹¹See Nelson, Beebout, and Skidmore (1985) for a summary of these suggested improvements.

The second approach--the price-variation approach--was suggested by the 1982 President's Private Sector Survey on Cost Control (the Grace Commission). It involves estimating the differences in prices per unit paid for food items by different-sized households. To account for the differences in the quality of food purchased by different-sized households, it is important to compare prices of food items that are similar in quality. A major disadvantage of estimating economies of scale with the price-variation approach is that it does not account for economies of scale associated with a more efficient use of food by larger households through such factors as reduced waste in food preparation and reduced spoilage in food storage.

b. Data Requirements

With either approach, a large sample of nationally representative households should be used to estimate the economies-of-scale adjustments used in the FSP. And as economies of scale may vary by income if, for example, higher income households have better access to stores that sell food in bulk, these adjustments should be estimated on a sample of low-income households. Kerr and Peterkin used a nationally representative sample of 4,376 nonfarm households. However, their sample contained only households with incomes above the poverty threshold and, thus, no low-income households.¹²

The expenditure-modeling approach to estimating economies of scale requires data on both the money value and quantity and types of food used by the household. The money value of the food used is needed to calculate per person food "expenditures;" the quantity and types of food used is needed to calculate the nutritional quality of the food. Other required data include household composition, household income, the value of any food stamps received by the household, and the number of meals eaten away from home by household members.

¹²The disadvantage of using a low-income sample for the expenditure-modeling approach is that the FSP benefit levels themselves may have an important effect on food expenditures. To control for this, the amount of food stamp benefit received should be included in the model.

The price-variation approach requires information on prices paid for food by different-sized households. To control for variation in the quality of food, it is important that the data provide detailed information on the specific types of food used.

3. Changing the FSP Benefit Issuance System, Criteria for Authorizing Retail Stores, and Coupon Denominations

Potential design changes to the FSP include changes in (1) the system used to issue food stamp benefits, (2) the criteria for authorizing retail stores to accept food stamps, and (3) the denomination of food stamp coupons. Making policy changes related to these aspects of program design requires data on food purchases made by food stamp recipients. While FNS proposed using the 1987-88 NFCS in research projects that would address changes to the TFP and economies-of-scale adjustments, FNS did not propose using the survey to address changes in the above three areas. However, because they are important aspects of the design of the FSP, we discuss the data that would be needed to make decisions about changes in these areas.

a. Form of Benefit Issuance

There is currently some policy interest in alternatives to the coupon form of the food stamp benefit. Two candidates are (1) cashout, in which recipients are provided with a check similar to an AFDC grant check, and (2) electronic benefit transfer (EBT), in which recipients are given a card similar to an automated-teller-machine (ATM) card that can be used to make food purchases at the point of sale by deducting the purchase amount from the recipient's food stamp account. These two alternative systems have been implemented in a number of demonstration sites.

It has been hypothesized that food stamp recipients who receive benefits in an EBT system will spend a greater proportion of their benefits on food because there are fewer opportunities to spend benefits on nonfood items by, for example, trafficking food stamp coupons for cash. A second hypothesis is that because of the enhanced security of benefits in an EBT system and the reduced stigma involved in purchasing food, recipients are more likely to spread their food purchases

throughout the month and are less likely to run out of benefits before they receive the next benefit. To evaluate alternative benefit issuance systems and test these hypotheses requires data on the timing, quantity, and types of food purchased by food stamp recipients.

b. Retail Store Authorization Policy

To be authorized to accept food stamps, a store is required to show to FNS that its participation in the FSP would "further the purposes" of the program. In practice, most stores that apply to accept food stamps are authorized to do so by FNS. However, it has been suggested that the authorization criteria be narrowed. This may improve the integrity of the program by reducing the number of stores that "traffic" in food stamps or accept food stamps in payment for unauthorized products. It would also free up the resources available to monitor the stores that do accept food stamps. The disadvantage of reducing the number of stores that accept food stamps is that food stamp recipients would have access to fewer stores that accept food stamps, which in turn interferes with recipient's ability to purchase food using food stamps. In order to assess its policy for authorizing retail stores to accept food stamps, FNS requires data on the quantity, nutritional content, and price of food purchased by food stamp recipients at different types of retail stores.

c. Coupon Denomination

Food stamp coupons are issued in booklets that contain coupons of \$1, \$5, \$10, and \$20 denominations. When using food stamps, recipients can only receive change of less than \$1 in the form of cash. If change of more than \$1 is required, the change can be provided in the form of coupons. In order to assess the most convenient and cost-effective denominations of coupons, FNS requires the following kinds of data on the purchases of food stamp recipients: how much recipients typically pay when they use food stamps, the form in which they pay (all coupons, or some coupons and some cash), whether they receive change, the number of times each month that purchases are made with coupons, and the types of stores in which these purchases are made.

C. MODELING PROGRAM PARTICIPATION

Models of program participation are often used to inform policymakers of the potential impact of proposed changes in legislation on both program cost and the number of people who receive benefits. Therefore, databases that provide information about program participation are a major research need for FNS. As in other research areas, the types of data needed to model participation in food assistance programs vary by program. However, in most instances, any database must meet three requirements if it is to be used to model program participation.

1. It should be based on a nationally representative sample containing households (or individuals, depending on the program) that are eligible for the program.
2. It should contain information on whether the household (or individual) participates in the program of interest.
3. It should contain sufficiently detailed information about the characteristics of the households (or individuals) to approximate their eligibility for the program.

In this section, we discuss a specific example of modeling program participation--updating the models of participation in the National School Lunch Program (NSLP) and the School Breakfast Program (SBP).

1. Modeling Participation in the NSLP and SBP

FNS uses models of participation in the NSLP and SBP to simulate the impact on program participation of program changes such as raising the income-eligibility criteria or increasing the reimbursements paid to schools for the cost of the meals. FNS currently uses a model of program participation based on data collected in 1983 and 1984 as part of the second National Evaluation of School Nutrition Programs (NESNP-2). Prior to the publication of the response rate to the 1987-88 NFCS, FNS proposed that the specification of the model should be revised and the model reestimated using data from the 1987-88 NFCS.

Income is not the only criterion that determines whether a child can receive free or reduced-price school meals. To participate in the school nutrition programs (1) the child must also attend a school that has elected to participate in the programs by providing USDA-approved lunches and/or breakfasts, (2) a parent or guardian of the child must have submitted an application for the child to receive free or reduced-price school meals, and (3) he or she must actually eat a meal at school. Therefore, modeling participation in school nutrition programs means modeling three decisions, or criteria that determine participation in the programs:

1. The school's decision to participate in the NSLP and/or SBP by providing meals
2. The decision by a parent or guardian to apply for the child to receive free or reduced-price meals
3. The child's decision to eat a school meal on any given day or on any day during a given period

Estimating models of participation in school nutrition programs requires a sample of school-aged children from households whose income does not exceed 185 percent of poverty. In order to be able to use the model to simulate the impact of changes in the program on total program costs and participation, the sample should be nationally representative. Ideally, the database should contain the following information on each child:

- Whether the child attends a school that participates in the school lunch and/or breakfast programs
- How frequently a child eats school lunches or breakfasts at school
- The full and reduced prices of the school lunches or breakfasts
- The child's "meal-price status," that is, whether the child can receive free or reduced-price meals or must pay full price for a school meal. Meal-price status cannot be inferred directly from whether the child is income-eligible for the program because although a child can be income-eligible for free or reduced-price school meals, if his or her parent has not applied to participate in the program, he or she will not be able to receive a free or reduced-price meal at school.

- Characteristics of the child and his or her family, including age, sex, grade level, race, ethnicity, household income, and participation in other assistance programs

D. DESCRIPTIVE STUDIES OF FOOD CONSUMPTION

In designing and administering its food assistance programs, FNS needs to understand the characteristics of the population it serves--low-income households and individuals. One of the more important characteristics of the low-income population is the nutritional quality of its diet. To keep informed about the dietary status of this population, FNS conducts descriptive studies of the food consumption of the low-income population. This section presents two examples of such studies. The first describes the food choices of low-income households. The second describes the food intake and other characteristics of low-income children. Before the release of the 1987-88 NFCS nonresponse rates, FNS planned to use the data from this survey for these and similar descriptive studies.

1. Describing the Food Choices of Low-Income Households

Kisker and Devaney (1988) used the 1979-80 SFC-LI to describe the food choices of low-income households. In order to investigate whether these choices have changed over the ten years since the fielding of the 1979-80 SFC-LI, FNS proposed using the low-income sample of the 1987-88 NFCS to update that study.

The proposed study was to address five research questions:

1. What percentage of low-income households use foods that meet the RDA for various nutrients?
2. Do the food choices of low-income households vary by such household characteristics as race, ethnicity, income, and employment status of the household head?
3. Do the food choices of low-income households vary by participation in food assistance programs?
4. Have the food choices of low-income households changed over time?
5. Are variations across subgroups of households attributable to differences in the quantity of food consumed or the types of food consumed?

Answering these questions requires a database with a large sample of low-income households. The 1979-80 SFC-LI, used by Kisker and Devaney, contains about 3,000 households that were eligible to participate in the FSP. The database also should contain information on household food use, nutrient availability, and participation in food assistance programs.

2. Describing the Food Intake of Low-Income Children

Because children comprise a large proportion of the low-income population, FNS is interested in understanding the characteristics of low-income children and the extent to which FNS programs meet their nutritional needs. A descriptive study is being conducted of existing food assistance programs and their compatibility with the nutritional needs and preferences of low-income children (defined as persons 18 years of age or younger) and their households. Areas examined in the study include the demographic and socioeconomic characteristics of low-income children, their participation in food assistance programs, and their intake of food and nutrients. In addition, the study is comparing the food and nutrient intake for the following groups of children:

1. Low-income children and children of all income levels
2. Children who participate in food assistance programs and children who are eligible for, but do not participate in, food assistance programs

The database for this study must contain a sample of low-income and higher-income children (the higher-income children must be included in the sample in order to make the comparison in item 1, above) in addition to information on the nutrient intake of children, participation in food assistance programs, and demographic and socioeconomic characteristics of their households. In its plans for this study, FNS indicated that the 1987-88 NFCS might be a useful source of data to describe the food intake of low-income children.

E. METHODOLOGICAL STUDIES

In addition to research used for the design and evaluation of its programs, FNS also sponsors research aimed at improving research techniques and methodologies. This section describes three examples of such projects. The goal of the first project is to develop a technique to adjust the distribution of average dietary intake for the day-to-day variation in a person's diet when only one day of intake is observed. The adjusted distribution would provide a better estimate of the distribution of usual dietary intake than does the unadjusted distribution. The goal of the second project is to develop a technique to address the problem of self-selection in models of multiple program participation. The goal of the third project is to test the validity of data collected in the evaluations of the Food Stamp Cashout Demonstrations. These research projects have very different data needs, but it was expected that each would use data from the 1987-88 NFCS.

1. Estimating the Distribution of Usual Dietary Intake with Only One Day of Dietary Intake Data

The measure of dietary intake that, ideally, should be used in assessing the dietary adequacy of a population is the usual daily intake, or the average daily intake that would persist over time. As most food consumption surveys collect data on individual dietary intake for no more than three days, usual daily intake is rarely measured directly. Instead, the observed dietary intake on one day or the average of daily intake over several days is typically used as a proxy for usual dietary intake.

The mean of the distribution of the average observed daily intake is an unbiased estimator of the mean of the distribution of usual dietary intake. However, as intake varies greatly from day to day, the variation across persons in *average* daily dietary intake is larger than the variation in *usual* dietary intake. Consequently, the distribution of average dietary intake is flatter (that is, it has a larger variance) than the distribution of usual dietary intake. This difference in the shape of the distribution is a serious problem because it can lead to bias in estimates of the proportion of the

population who are nutritionally at risk if those estimates are based on the distribution of average dietary intake.

To correct this problem, the Subcommittee on Criteria for Dietary Evaluation of the National Research Council (NRC, 1986) proposed a statistical procedure for correcting the bias in the variance of the distribution of average dietary intake. The procedure entails estimating an "adjustment factor" that is used to pull each point in the distribution of the average dietary intake toward the mean. The adjustment factor is calculated from a measure of the intraindividual variation, or the variation from day to day in each person's dietary intake. Under certain assumptions specified by the NRC (1986), the variance of the adjusted distribution of average dietary intake is an unbiased estimate of the variance of the distribution of usual dietary intake. At least two days of dietary intake data from some of the individuals are required to estimate the adjustment factor.

Many food consumption surveys collect only *one* day of dietary intake data from each person in the sample. The distribution of observed one-day dietary intake calculated from these databases will be flatter than the distribution of usual dietary intake, but the sample cannot be used to estimate an adjustment factor to correct for the bias in the variance of the distribution because only one day of intake data are available. One proposed solution to this problem is to obtain an estimate of the adjustment factor from a survey that includes multiple days of intake data and to apply this adjustment factor to the distribution of average dietary intake estimated from the sample of one-day intakes. Burghardt et al. (1990) suggested using the multiple days of intake data from the 1987-88 NFCS to estimate an adjustment factor and applying this estimated adjustment factor to the distribution of average dietary intake of school children in the School Nutrition Dietary Assessment study.

In their analysis of the CSFII, Ritenbaugh et al. (1988) found that the degree of intraindividual variation, and hence the adjustment factor, varied substantially across population strata. Hence, if an adjustment factor estimated from one sample is to be used to adjust the intake distribution

estimated from another sample, it is important that the two samples are drawn from the same population during roughly the same time period. For example, an adjustment factor estimated from a sample of children should not be used to adjust the intake distribution of adults.

2. Developing Methodologies for Analyzing Self-Selection in Models of Multiple Program Participation

"Self-selection" is often a problem in analyzing food assistance programs. It occurs when program participants differ from nonparticipants in ways that are related to the outcome of interest but are not measured in the data, thus creating a bias in conventional estimates of program effects. To reduce self-selection bias in estimates of program effects, analysts typically use a variant of a self-selection adjustment technique developed by Heckman (1978, 1979) and Heckman and Robb (1985).

Many recipients of food assistance receive benefits from more than one program, and they often receive benefits from three or more programs. For example, low-income school children may participate in the SBP, the NSLP, and the FSP. Estimating program models that incorporate the possibility of participation in three or more programs and correcting for self-selection bias is extremely difficult using the standard Heckman techniques. Even with only two programs, the econometric software that has been developed to estimate the models is difficult to use. Estimating models of participation in three or more programs would be very expensive and would require considerable mainframe computer resources.

However, McFadden (1986) and Pakes and Pollard (1986) developed a new technique--the method of simulated moments--that can be used to estimate models of multiple program participation in the presence of self-selection. The advantages of this technique are that it is relatively simple to implement, it requires fewer computer resources than the Heckman technique, and it can be used to model participation in any number of programs. FNS wished to develop an econometric model and the accompanying computer software to estimate the impact of multiple program participation

on household food expenditures and nutrient availability using this technique. It planned to test the econometric model and computer software on data from the 1987-88 NFCS.

In this research project, the estimates derived from the models are less important than the development of the techniques used to make the estimates. Hence, the main requirements of the database used for this type of research are that it contain a large enough sample and the correct variables. It is not so important that the sample be representative of all low-income U.S. households.

3. Validating Data From the Food Stamp Cashout Demonstration Evaluations

FNS is currently sponsoring three evaluations of food stamp cashout demonstrations in which the principal outcome of interest is the impact of cashout on household food consumption. These evaluations, which are being conducted in Washington State, Alabama, and San Diego County, entail the collection of data on household food expenditures and food use from samples of food stamp recipients. Each of the three independent samples from each area is divided approximately equally between recipients of food stamp coupons and recipients of food stamp checks. The cashout evaluations will compare the average values of food expenditure, food use, and nutrient availability of check households to those of coupon households. When differences are found, statistical tests will be conducted to determine whether the data can support a finding that cashout altered the food consumption behavior of food stamp recipients in a particular demonstration site.

A major limitation of the cashout evaluations is that they are based on samples drawn from three very limited geographic areas. These samples were selected to be representative of the food stamp caseloads in the demonstration areas; they are not representative of the full national caseload of food stamp recipients. While it will therefore not be possible to formally generalize from the findings based on the cashout samples to the national caseload, FNS will need to develop some tentative conclusions about the effects of implementing cashout nationwide. The development of those conclusions could be facilitated by first conducting two types of comparisons between food stamp recipients in the cashout demonstration sites and food stamp recipients nationwide. The first

comparison would focus on household economic and sociodemographic characteristics. The second would be a comparison of food consumption. The latter would compare recipients of food *coupons* in the demonstration sites to recipients of food coupons nationwide. These comparisons would provide estimates of the extent to which the food stamp caseloads in the demonstration sites resemble the food stamp caseload nationwide. These comparisons could guide FNS in drawing tentative conclusions from the cashout evaluations about the effects of cashout if it were adopted on a nationwide basis.

III. THE 1987-88 NATIONWIDE FOOD CONSUMPTION SURVEY (NFCS)

The 1987-88 NFCS is the seventh in a series of national food consumption surveys conducted by HNIS approximately every ten years. The first four surveys collected information only about food used by households; the last three surveys, conducted since 1965, have also collected individual food intake data. The 1987-88 NFCS was designed to collect information about the types, nutrient content, and money value of food used in U.S. households and about the food intake of individuals while at home and away from home. These data are used for a wide variety of purposes including policy formulation, regulation, nutrition education, research, and design and evaluation of food assistance programs.

This chapter describes the 1987-88 NFCS and discusses the potential problems of using the survey data. Section A describes the design and content of the survey. Section B discusses potential problems with the survey and recent reviews of the survey by HNIS and the LSRO expert panel. Section C discusses the advantages and disadvantages of merging the basic and low-income samples of the 1987-88 NFCS. Section D summarizes the extent to which the 1987-88 NFCS provides useful food consumption data.

A. DESCRIPTION OF THE 1987-88 NFCS

The NFCS consists of two samples: a basic sample and a low-income sample. Both are stratified area probability samples of private households in the 48 contiguous states and all individuals who live in those households. The basic sample includes households with all levels of income. The low-income sample includes only households that in the previous month had a total income (excluding food stamps and WIC benefits) of 130 percent of the poverty level or less.¹

¹For some of these households, income during the previous year exceeded 130 percent of poverty.

The survey was originally planned to be implemented for 12 months beginning in April 1987. However, because of problems in interviewing households, the period was extended by five months and ended in August 1988. Data on both households and individuals from the basic sample are now publicly available. Data on households from the low-income sample will be publicly available in the summer of 1992. Data on individuals data from the low-income sample will not be released either publicly or internally within USDA.

The 1987-88 NFCS differs from other recent national food consumption surveys in that it collected information on both food used at home by the *household* and on *individual* food intake at home and away from home for all household members. Data on household *nutrient* availability and individual *nutrient* intake are derived from the quantities of food used at home and individual food intake using the HNIS National Nutrient Data Bank (Hepburn, 1982). The sample of households is the same for both the household-level data and the individual-level data, but the methodologies for collecting the two types of data differ. The kinds of data that were collected, along with the methods of data collection, are discussed below.

1. Household Food Use Data

The 1987-88 NFCS collected data on all food used at home during a seven-day period. It was collected by an aided recall during an in-person interview with the household's main meal planner or preparer. The respondent was contacted at least seven days before the interview and asked to keep grocery receipts, food labels, recipes, and other reminders of the food that would be used during the week before the interview. During the interview, the respondent was asked whether the household had used any food in the previous seven days from each of 21 food groups. If the respondent indicated that the household had used food from a particular group, the interviewer then asked for more specific information about foods used within that group, including the type and quantity of food consumed; the form of the food (for example, fresh, frozen, or canned); whether it was bought, received as a gift, or home produced; and if bought, the quantity and unit price of the

food. The respondent was also asked to recall the household's usual expenditure on food for home consumption per week or per month over the past three months.

2. Individual Food Intake Data

The survey requested food intake data for three consecutive days from each individual in the household. Each individual who was 12 years of age or older and at home at the time of the interview was interviewed separately. The main meal planner or preparer responded for children younger than age 12 and, if possible, for absent household members. Each respondent was asked to recall all foods eaten on the previous day both at home and away from home. Each respondent was also asked to maintain a record of food consumption for both the day of the interview and the next day. Information was collected on the type and quantity of food consumed as well as when, with whom, at what meal the food was consumed, and whether it was eaten at home or away from home.

3. Household and Individual Characteristics

In addition to food consumption data, the 1987-88 NFCS collected information on household demographics, income over the past year, assets, and socioeconomic characteristics. The survey collected information about the age and sex of each member of the household and about whether any women in the household were pregnant or breastfeeding. The main meal planner or preparer was also asked about the household's participation in the following food assistance programs:

- ***Food Stamp Program (FSP).*** Information includes whether the household currently receives food stamps or received food stamps in the previous 12 months, the value of the food stamps, and the date the household last received food stamps.
- ***Special Supplemental Food Program for Women, Infants, and Children (WIC).*** Information includes who, if anyone, in the household received WIC benefits in the previous month.
- ***National School Lunch Program (NSLP) and School Breakfast Program (SBP).*** Information includes whether any child in the household attends a school that serves lunches or breakfasts, how frequently the child receives a lunch or breakfast, and how much the child pays for the meals.

- ***Receipt of USDA Surplus Commodities.*** Information includes whether the household received any cheese or butter in the previous three months, when it received the cheese or butter, and the quantity of cheese and butter received.

B. POTENTIAL PROBLEMS IN THE 1987-88 NFCS

There are two potential problems in the 1987-88 NFCS: a high rate of nonresponse and unequal coverage over the seasons of the year and the days of the week. Both problems were a function of survey implementation rather than its original sampling design.

Because of the high rate of nonresponse to the 1987-88 NFCS, the actual sizes of the basic and low-income samples of the survey were much lower than planned. In the basic sample, only 38 percent of eligible households participated in the survey. Of the individuals living in the *participating* households, 81 percent provided at least one day of intake data, and 68 percent provided all three days of intake data. Of the estimated number of individuals living in *eligible* households, only about 31 percent provided one day of intake data, and only about 26 percent provided all three days of intake data. These response rates are much lower than response rates for such similar surveys as the Continuing Survey of Food Intakes by Individuals (CSFII). In the 1986 CSFII, 79 percent of eligible households participated in the basic survey, and 88 percent of eligible households participated in the low-income survey. The nonresponse rates for the low-income sample of the 1987-88 NFCS are not yet available; however, the survey was designed to collect information from about 8,000 low-income individuals, and only about 6,400 such individuals provided at least one day of intake data.

The low response rate has two implications for estimates based on the 1987-88 NFCS. First, because the sample size is smaller than planned, estimates will be less *precise*. By less precise, we mean that the estimates will have higher variances and standard errors. Smaller sample sizes also make it more difficult, and perhaps even unfeasible, to analyze subgroups of the population. Second, if respondents differ from nonrespondents in their food consumption, estimates based on the sample will be *biased*. A biased estimate of a variable systematically underestimates or overestimates the true value of the variable.

The design of the 1987-88 NFCS specified that the household samples selected and interviewed each quarter would all be the same size. However, because of the ways in which the data were collected, an unequal number of households was sampled each quarter. As a result, the observations are not equally distributed across seasons: 33 percent of all households in the basic survey were interviewed in the fall; 26 percent in the winter; 22 percent in the summer; and 18 percent in the spring. Similarly, the observations are not equally distributed across the days of the week. Only 4.7 percent of the observations of individual food intake are for a Saturday because fewer interviews than planned were conducted on a Sunday. This figure is much lower than the 14 percent of observations that would have been made had the data been collected evenly throughout the week. If uncorrected through weighting procedures, this unequal coverage of respondents over seasons of the year and days of the week could lead to biased estimates. Below, we discuss in more detail the implications of nonresponse and unequal coverage.

1. Loss of Precision Resulting from Nonresponse

Nonresponse created smaller basic and low-income samples than were planned for the 1987-88 NFCS. Table III.1 presents the actual and target sample sizes of the 1987-88 NFCS. In the basic survey, only about 4,500 households, or 75 percent of the target number, completed the household interview. In the low-income survey, only about 2,500 households, or 70 percent of the target number, completed the survey. Similarly, only about 10,300 individuals in the basic survey, or 74 percent of the target number, provided at least one day of intake data; and in the low-income survey, only about 6,400 individuals, or 81 percent of the target number, provided at least one day of intake data.²

The smaller samples lead to a loss of precision; that is, estimates made using the sample will have higher variances and standard errors. The loss of precision is especially important in analyzing

²The actual sample sizes are 70 to 81 percent of the target sample sizes despite much lower response rates because survey field staff attempted more interviews than had been originally planned.

TABLE III.1
ACTUAL AND TARGET SAMPLE SIZES OF THE 1987-88 NFCS

Unit of Observation	Actual Sample Sizes		Target Sample Sizes	
	Basic	Low-Income	Basic	Low-Income
Households	4,495	2,509	6,000	3,600
Individuals				
With at least 1 day of intake data	10,342	6,443 ^a	14,000	8,000
With at least 3 days of intake data	8,468	5,309 ^a	N/A	N/A

SOURCE: The actual number of households and individuals in the basic sample and the actual number of households in the low-income sample were counted from the 1987-88 NFCS (unweighted). The number of individuals in the low-income sample is a preliminary count reported in Hubbard et al. (1991). The target sample sizes for both samples are reported in Appendix A of the LSRO report (LSRO, 1991).

^aPreliminary count

N/A: not available

subgroups of the population because many of the applications of the 1987-88 NFCS require an analysis of subgroups that while small, are important to policymakers. For example, an evaluation of the WIC program would involve an analysis of the subgroup comprising pregnant, breastfeeding, and postpartum women and children younger than age 5 in households with income of less than 185 percent of poverty (that is, those individuals who may be eligible for WIC benefits).

Tables III.2 and III.3 show the number of households and individuals, respectively, in the 1987-88 NFCS in some subgroups of interest to FNS. These tables give a good idea of just how few members of certain subgroups are actually represented in the survey. For instance, only about 1,000 households in the basic sample and about 2,000 households in the low-income sample have annual income no greater than 130 percent of poverty. Among the households in the basic sample with income no greater than 130 percent of poverty, only about 2,500 individuals provided at least one day of food intake data, and fewer than 2,000 individuals provided three days of food intake data. Only about 530 households in the basic sample and 800 households in the low-income sample have a black household head. There are only 62 pregnant or breastfeeding women who are income-eligible for WIC in the basic sample. The sample sizes for some of these subgroups are so small that it would be impossible to conduct meaningful analyses of the subgroups using the 1987-88 NFCS. For most of these subgroups, sample sizes would have been problematic even if the overall response rate to the 1987-88 NFCS had been high. Either a larger overall target sample size or oversampling of these subgroups would have been necessary to achieve adequate sample sizes for these subgroups.

2. Bias Resulting from Nonresponse and Incomplete Coverage

If household or individual respondents differ from nonrespondents in food consumption (or other outcomes of interest), estimates based on the sample will be biased. This phenomenon is known as *nonresponse bias*. Nonresponse bias occurs because the respondents are not truly representative of the population. For example, suppose that individuals in low-income households are more likely to participate in the survey and also consume less protein than individuals in non-low-income

TABLE III.2

NUMBER OF HOUSEHOLDS IN THE BASIC AND LOW-INCOME SAMPLES OF THE
1987-88 NFCS BY CHARACTERISTIC OF HOUSEHOLD

Household Characteristic	Number of Households (Unweighted)	
	Basic Sample	Low-Income Sample
Household had annual income last year no greater than 130 percent of poverty	1,010	2,089
Household had annual income last year no greater than 185 percent of poverty	1,546	2,361
Household is currently receiving food stamps	314	838
Household received WIC benefits last month	151	284
Head of household is black	531	808
Sample Size	4,495	2,509

SOURCE: Data are derived from the 1987-88 NFCS (unweighted).

TABLE III.3

NUMBER OF INDIVIDUALS IN THE BASIC SAMPLE OF THE 1987-88 NFCS BY
CHARACTERISTIC OF INDIVIDUAL

Individual Characteristic	Number of Individuals (Unweighted)	
	One Day of Intake Data	Three Days of Intake Data
Individual is in household with annual income no greater than 130 percent of poverty	2,485	1,955
Individual is in household with annual income no greater than 185 percent of poverty	3,777	3,034
Individual is in household currently receiving food stamps	859	666
Individual receives WIC benefits	208	173
Child under 5, in a household with annual income no greater than 185 percent of poverty	417	339
Pregnant or breastfeeding woman, in household with annual income no greater than 185 percent of poverty	62	49
Child attends a school that serves school lunches and is in a household with annual income no greater than 185 percent of poverty	842	690
Child attends a school that serves school breakfasts and is in a household with annual income no greater than 185 percent of poverty	343	272
Individual is black	1,181	905
Sample Size	10,342	8,468

SOURCE: Data are derived from the 1987-88 NFCS (unweighted).

households. As a result of the fact that individuals from low-income households would be overrepresented in the sample, the mean intake of protein for the total population would be underestimated.

To investigate the differences between respondents and nonrespondents, HNIS compared 13 demographic characteristics of households responding to the 1987-88 NFCS with the characteristics of households in the March 1987 Current Population Survey (CPS). As the CPS has a high response rate, it is assumed that it represents the true population of households. The comparisons were made using three age-sex subsamples: (1) men 20 years of age and older, (2) women 20 years of age and older, and (3) persons younger than 20 years of age. Using the basic sample of individuals who provided at least one day of intake data, six differences between the types of households responding to the 1987-88 NFCS and the types of households in the CPS were found to be statistically significant:³

1. The 1987-88 NFCS basic sample contains a larger proportion of individuals from low-income households than the March 1987 CPS.
2. The 1987-88 NFCS basic sample contains a larger proportion of individuals from two-adult households than the March 1987 CPS.
3. The 1987-88 NFCS basic sample contains a smaller proportion of women from households with working female household heads than the March 1987 CPS.
4. The 1987-88 NFCS basic sample contains a smaller proportion of men from households with working male household heads than the March 1987 CPS.
5. The 1987-88 NFCS basic sample contains a smaller proportion of men and women 20 years of age and older from households with a female household head younger than 41 years of age and with no children younger than age 18 than the March 1987 CPS.
6. The 1987-88 NFCS basic sample contains a smaller proportion of persons between 15 and 24 years of age than the March 1987 CPS.

³These comparisons were made using unweighted data from the 1987-88 NFCS and weighted data from the March 1987 CPS.

These differences led the LSRO expert panel to conclude that the 1987-88 NFCS underrepresented "nontraditional families" (LSRO, 1991). The differences in the characteristics of respondents do not, by themselves, indicate the presence of nonresponse bias. However, if these differences are related to the outcome of interest, and the respondents in the March 1987 CPS represent the true population of interest, there will be nonresponse bias. The LSRO expert panel concluded that all six factors were likely to be correlated with food consumption.

Incomplete coverage of the seasons of the year and days of the week will also lead to bias in estimates based on the sample if respondents' food consumption varies by season or day of the week. Studies have shown that the intake of some nutrients is related to the seasons and that the intake of food energy is highest on weekends (Battesse et al., 1988). Thus, if procedures are not used to correct for incomplete coverage, estimates of mean intake based on the 1987-88 NFCS will be biased because observations in the spring and on Saturday are underrepresented, and observations in the fall and on weekdays are overrepresented.

a. Bias Corrected by Weighting

If respondents differ from nonrespondents in known and measurable ways, nonresponse bias can be corrected by weighting the data. Each observation is multiplied by a weight, where the sum of the weights is equal to the population total. Households or individuals from groups who are underrepresented in the sample are weighted by larger factors than are households or individuals from groups who are overrepresented. Bias caused by incomplete coverage by day and season can be similarly corrected by weighting procedures. Observations reported in seasons and days that are underrepresented in the sample receive a higher weight than those reported for more fully represented seasons or days.

HNIS, working with statisticians at Iowa State University (Loughin and Fuller, 1990), weighted the 1987-88 NFCS basic sample to correct for nonresponse and unequal coverage of seasons and days. A regression-based technique was used to estimate the weights. In constructing the weights

to be applied to the households, HNIS controlled for season of the year and the following 15 demographic characteristics: household size; the square of the household size; geographic region; age of household head; race of household head; degree of urbanization; household income below the poverty level; ownership of home; whether the household has a male head, a female head, or both; presence of a female head of household who worked in the previous week; presence of one adult in the household; presence of two adults in the household; presence of a child younger than 7 years of age; presence of a child between 7 and 17 years of age; and presence of a household member receiving food stamps. The "controls" used in constructing the individual weights were similar but not exactly the same. For example, the degree of urbanization and household size were not included as controls in calculating the individual weighting factors. The estimation procedure made each weight positive and the sum of the weights equal to the population total in thousands taken from the March 1987 CPS. The weights for the household component of the low-income sample will be included in the public use file for the household data, to be released in the summer of 1992.

While estimates based on weighted data may be less biased, they are also less precise than those based on unweighted data. Hence, weighting the data exacerbates the loss of precision already associated with the small sample size created by nonresponse. The greater the range of weights used, the greater the loss of precision. In the 1987-88 NFCS basic sample, the weights vary considerably. While the average individual weight is 23.5 (that is, on average, one individual in the sample represents 23,500 individuals in the population), the weights range from 1 to 78 for women 20 years of age and older, 1 to 130 for men 20 years of age and older, and 1 to 136 for persons younger than 20 years of age. The average household weight in the basic sample is 19.8, and those weights range from 6 to 47. The large range in the individual and household weights is a direct reflection of large differences in response rates across population subgroups.

When some observations in a sample are given large weights, estimates made from the weighted data can be sensitive to the exclusion of a few observations. For example, Devaney and Fraker

(1989) found that excluding 19 out of approximately 4,500 observations from the weighted low-income sample of the 1977-78 NFCS changed their estimate of the "marginal propensity to consume food" out of food stamps from 0.41 to 0.27. The 19 excluded households were those households with weights larger than 20; the average weight in the sample was just under 1.

Moreover, analysts who use weighted data cannot use standard statistical packages because, when weights are used, these packages yield biased estimates of variances and standard errors. However, such statistical packages as SUDAAN, OSIRIS, and PC/CARP correct the variances and standard errors for the impact of the weights. In addition, these packages correct the standard errors and variances for the biases that occur because the survey did not consist of a simple random sample but had a more complex stratified design.

b. Bias Not Corrected by Weighting

If respondents differ from nonrespondents in unknown or unmeasurable ways that are related to an outcome of interest, nonresponse bias cannot always be corrected by weighting. For instance, if factors not considered by the weighting procedures affected both food consumption and the decision to respond to the survey, estimates from the 1987-88 NFCS will still be biased even if weights are used. These factors may either be observed, but not used, in the weighting procedure or unobserved, and hence unusable, in the weighting procedure.

An example of an *observed* factor that may affect both the decision to respond to the survey and food consumption is the degree of urbanization. Due to confidentiality restrictions, this variable is suppressed on about 18 percent of the March 1987 CPS records that were used to compute the NFCS weights. Because of this, HNIS did not control for the degree of urbanization in estimating the individual weights for the 1987-88 NFCS. However, HNIS did control for urbanization in constructing the household weights because urbanization was believed to have a much stronger correlation with household food use than with individual food intake (LSRO 1991, Appendix J). The

urbanization weights were derived not from the incomplete CPS data but rather from the basic design of the NFCS household sample and its allocation across three urbanization categories.

The following are two examples of *unobserved* factors that may affect both the decision to respond to the NFCS and food consumption:

- **Concern about Nutrition.** Individuals concerned about nutrition may be more likely to respond to a food consumption survey and may also have higher quality diets than individuals who are not concerned about nutrition.
- **Value of Time.** Individuals who highly value their time may be less likely to respond to a food consumption survey and are likely to have diets that are different from individuals who place a lower value on their time.

Not only can weighting fail to correct for bias, under certain conditions it can actually exacerbate the nonresponse bias. This would be the case if unrepresentative individuals were given large weights. For example, if respondents in households with a working female head are given large weights to correct for nonresponse by this group, and these individuals are not representative of individuals in all households with a working female head, the bias caused by the difference between respondents and nonrespondents may be exacerbated by the large weights.

c. Determining the Existence of Bias

While the preceding examples illustrate nonresponse bias that may not be corrected by weighting, it is impossible to state whether such bias exists without the benefit of detailed information about the households and individuals that did not respond to the survey. Because such information does not exist, we will never be able to say whether serious nonresponse bias exists in the 1987-88 NFCS. Recognizing this limitation, HNIS worked with the LSRO expert panel to conduct studies to detect evidence of potential nonresponse bias. These studies involved (1) comparing the characteristics of individuals and households by their degree of response to the survey and (2) comparing nutrient intake observed in the 1987-88 NFCS with nutrient intake observed in the 1977-78 NFCS and the 1985-86 CSFII.

Although little is known about the characteristics of households that did not respond at all to the 1987-88 NFCS, more is known about the households that participated in, but did not complete, the survey. HNIS divided the households that participated in the basic survey into three groups according to the degree of response: (1) households that completed the household part of the survey but did not provide any individual food intake data, (2) households that completed the household part of the survey and provided one day of individual food intake data, and (3) households that completed the household part of the survey and provided two or three days of individual food intake data. HNIS compared the characteristics of the households across each of these three groups, the presumption being that factors that affect the degree of response to the survey are also likely to affect whether the household participated at all.

HNIS found that some factors that affect the degree of response also affect nutrient intake: race, urbanization, and income as a percent of poverty were found to be correlated with the degree of response and with the intake of food energy and fiber. For example, in the basic sample, 13 percent of all participating households provided no individual food intake data, 11 percent provided only one day of intake data, and 75 percent provided more than one day of intake data. In contrast, of all participating households with a black head, about 19 percent provided no individual food intake data, and only 67 percent provided more than one day of intake data.

If the decision to participate in the survey at all is related to the degree of response to the survey, these findings suggest that, at least in the unweighted data, there is nonresponse bias. However, race, urbanization, and income as a percent of poverty were all used as controls in constructing the household weights; race and income as a percent of poverty were also used as controls in constructing the individual weights. Thus, we cannot conclude that nonresponse bias is still present after weighting. However, the existence of observable factors that are correlated with both nutrient intake and the high incidence of nonresponse to the survey highlight the concern that there may be other unobservable factors, not used in the construction of the weights, that are

correlated with both nutrient intake and the degree of response. If this is the case, there would be nonresponse bias even after weighting.

HNIS also compared the mean intakes of poultry, milk, fruit, food energy, and protein in the 1987-88 NFCS with the corresponding mean intakes in the 1977-78 NFCS and in the 1985 and 1986 CSFIs. These comparisons were made using weighted data. While no significant trend was found to be associated with the consumption of poultry, milk, fruit, or protein, the intake of food energy was found to be significantly lower in the 1987-88 NFCS than in the 1985 and 1986 CSFIs. One interpretation of this apparent difference in food energy intake is that it is purely a result of nonresponse bias in the 1987-88 NFCS. If respondents to the 1987-88 NFCS consumed less food energy than nonrespondents, estimates of the average intake of food energy based on survey data would be low. However, changes in the design of the survey, in the translation of food into its component nutrients, or even real changes in consumption patterns could also explain this trend.

C. COMBINING DATA FROM THE BASIC AND LOW-INCOME SAMPLES OF THE 1987-88 NFCS

We have identified two potential problems with the 1987-88 NFCS that are associated with the high rate of nonresponse: (1) because of the reduced sizes of the basic and low-income samples, estimates based on the 1987-88 NFCS are less precise than anticipated, and the number of observations on certain subgroups are too small to support planned analyses; (2) if respondents to the survey differ from nonrespondents in ways that are related to outcomes of interest, estimates of those outcomes based on the survey data may be biased. In Section B, we explained that HNIS has taken steps to reduce the problem of biased estimates by weighting the NFCS data. Unfortunately, there is reason to believe that some bias remains despite the weighting, and the extent of that bias is unknown. This section discusses a potential solution to the problem of imprecise estimates and small numbers of observations on population subgroups--combining the data from the survey's basic and low-income samples.

1. The Potential Benefits of Combining Data From the Two Samples of the 1987-88 NFCS

At least twice in recent years, FNS has sponsored efforts which combined data from the basic and low-income samples of USDA food consumption surveys.⁴ The objective of combining data from the basic and low-income samples of the 1987-88 NFCS would be the same as the objective for combining data from other surveys--to generate a data file that would contain more observations on low-income households and their members than are available in either the NFCS basic or low-income files alone. Since approximately 1,000 households in the 1987-88 NFCS basic sample have incomes no greater than 130 percent of poverty,⁵ the combining of data from the two samples would result in a single data file that contains about 3,500 observations on low-income households, a 40 percent increase over the 2,509 households actually in the low-income sample. The combined file would support more precise estimates of diet-related and other outcomes for low-income households and their individual members. In addition, because it would provide substantially greater numbers of observations on subgroups of the low-income populations, the combined data file would support analyses of those subgroups that might not otherwise be feasible.

2. The General Problem of Combining Data From a Basic Sample and a Low-Income Sample

Under the best of circumstances, combining data from a basic sample and a low-income sample can be challenging and subject to error. At issue is the problem associated with combining data from a sample into which households were selected without regard for local poverty rates (basic sample) with data from a sample into which households in high-poverty areas were selected with a greater

⁴Under contract to FNS, MPR combined data from the basic and low-income samples from the first wave of the 1985 CSFII (Fraker and Post, 1987) and from the first wave plus three randomly selected subsequent waves of that survey (Fraker and Post, 1989).

⁵The estimate of 1,000 low-income households in the NFCS basic sample is based upon an annual measure of income (refer back to Table III.2). Given that households were screened into the low-income sample on the basis of their income during the month that preceded the interview, it would be preferable to identify low-income households in the basic sample on the basis of an equivalent measure of monthly income. If we did this, we would expect to find somewhat more than 1,000 low-income households in the basic sample.

probability than were households in low-poverty areas (low-income sample). In general, estimates based on a given number of cases that were selected with equal probability have smaller standard errors than estimates based on the same number of cases that were selected with unequal probabilities. After combining data from a basic and a low-income sample, the data must be reweighted to reflect two conditions: (1) low-income households are overrepresented in the combined sample, and (2) because of the reduced precision of estimates based on households from the low-income sample, less weight should be placed on those cases in computing an estimate.

As previously noted, MPR has twice combined data from the basic and low-income samples of USDA food consumption surveys. Each of these tasks required substantial input from a sampling statistician and other high-level technical staff over a period of several months. Consequently, the tasks were relatively expensive to complete.

3. The Specific Problem of Combining Data From the Two Samples of the 1987-88 NFCS

The high rate of nonresponse to the 1987-88 NFCS substantially complicates the general problem of combining data from samples with different designs in two ways: (1) the weights for the combined data must incorporate both the weights that were already computed by HNIS and Iowa State University for the basic sample and the corresponding weights for the household component of the low-income sample;⁶ (2) any nonresponse bias that remains in estimates generated from the separate samples when the HNIS/Iowa State weights are used will also be present in estimates based on the combined data. The first complication implies that the combining and weighting of data from the basic and low-income samples of the 1987-88 NFCS would be more challenging and more subject to error than is the case when nonresponse is not of concern. The second complication implies that even if the substantial task of combining and weighting the data from the two samples were

⁶As previously noted, HNIS will not be computing weights for the individual component of the low-income sample, nor will it be releasing those data. Thus, this discussion of combining data from the two sample of the 1987-88 NFCS applies only to the household components of those samples.

undertaken, any concern about nonresponse bias in estimates based on either sample alone would exist to at least the same degree in estimates based on the combined sample.

Simply stated, sample nonresponse complicates the already challenging task of combining of data from samples with different designs, and while combining data increases the effective sample size, thus reducing the standard errors of estimates, it does nothing to reduce the problem of nonresponse bias. Consequently, combining the household data from the basic and low-income samples of the 1987-88 NFCS would make sense only if the benefits of (1) estimates with smaller standard errors and/or (2) more observations on subgroups outweigh the cost of the effort to combine the data. In making this assessment (or in deciding whether this trade-off is worthwhile), it must be recognized that estimates based on the combined data may still be biased as a result of nonresponse, and therefore vulnerable to challenges about their reliability.

D. USEFULNESS OF THE 1987-88 NFCS

The 1987-88 NFCS has at least five characteristics that, in principle, make the data particularly useful for FNS. First, it contains a large sample of low-income households. The low-income sample is especially important for studies of food assistance programs. Even though the actual number of households in the 1987-88 NFCS is much lower than planned, the low-income sample still contains over 2,500 households, and the basic sample still contains over 1,000 households with income at or below 130 percent of poverty. The size of the low-income sample of the 1987-88 NFCS may still be large enough for some applications.

Second, the NFCS is the only major survey that collects detailed information on household food use. This is important because household food use is one of the main outcomes of interest in evaluations of the impact of the FSP.

Third, the NFCS collects information on three measures of food consumption from each household: (1) household food expenditure, (2) household food use, and (3) individual nutrient

intake. These are the three types of food consumption data that FNS requires to address its full research agenda.

Fourth, the NFCS provides food intake data for more than one day. Estimates of the distribution of usual nutrient intake are required for any evaluation of the impact of food assistance programs on the proportion of households that are at nutritional risk. At least two days of intake data are required from at least some individuals in the sample to estimate this distribution. The fact that the NFCS provides up to three days of intake data makes the survey particularly useful.

Finally, the NFCS contains a wealth of information about participation in food assistance programs, including the FSP, WIC, NSLP, and SBP.

In practice, the utility of the 1987-88 NFCS to FNS in carrying out its research agenda is limited by the high rate of nonresponse to the survey. At best, nonresponse leads to smaller-than-planned sample sizes, which makes estimates less precise and subgroup analyses more difficult to conduct. Although the effective sample sizes could be increased by merging the low-income and basic samples, this process would be potentially costly and especially complicated given the high and variable nonresponse rates.

At worst, nonresponse can also potentially lead to seriously biased, and hence misleading, estimates. To some extent, this bias can be corrected by weighting the data, although developing appropriate weighting procedures is not a simple process. If respondents differ from nonrespondents in ways that are not captured by the weights (as is most likely the case), weighting the data will not fully correct the bias, and may even exacerbate it. In addition, the weighting procedures can lead to a further loss in precision.

In any survey without a 100 percent response rate, nonresponse bias is a potential problem. The higher the rate of nonresponse, the greater the potential problem of nonresponse bias. Without data on food consumption from households that did not respond to the 1987-88 NFCS, we cannot identify the extent of nonresponse bias. However, the extremely high rate of nonresponse in the survey does

tell us that the problem of nonresponse bias is potentially serious. In fact, the LSRO expert panel judged the potential problems from nonresponse as serious enough to state that they "do not recommend use of the data from the 1987-88 NFCS" (LSRO 1991, 13).

IV. ALTERNATIVE SOURCES OF DATA ON FOOD CONSUMPTION

Many of the research projects conducted by FNS require data on both food consumption and participation in USDA food assistance programs. FNS has previously relied on the NFCS to satisfy many of its needs for such data. However, given the potential problems with the 1987-88 NFCS, there is some doubt about whether FNS should use it in its planned research projects. If it is not used, alternative sources of data need to be found to meet FNS' research needs. This chapter describes databases that are potential substitutes for the 1987-88 NFCS.

Most of the databases discussed in this chapter are generated by large-scale surveys conducted at regular intervals. However, we also discuss some databases, such as the Food Stamp Cashout Evaluation databases, that were or will be created as part of special one-time research projects. In addition, we discuss other databases derived from retail sales data (retail food store scanner data) that are created on a weekly or monthly basis. The possibility of conducting a new survey to collect data to meet FNS' research needs is discussed in Chapter VI.

Although FNS' specific needs for food consumption data vary, for the purposes of most applications the database should fulfill five general requirements:

1. It should contain data on food expenditure, household food use, and individual food intake.
2. It should contain information about eligibility for and participation in USDA food assistance programs.
3. It should be large enough to yield precise estimates and to allow for some subgroup analyses.
4. It should contain information on sociodemographic and economic characteristics of the household.
5. It should be fairly current.

It is unfortunate that no existing database other than the NFCS satisfies all five requirements. Table IV.1 summarizes the important characteristics of the 1987-88 NFCS and the other databases described in this chapter. The table highlights the importance of the 1987-88 NFCS. No other database except for its precursors, the 1977-78 NFCS and the 1979-80 Survey of Food Consumption in Low-Income Households (SFC-LI), contains information on household food expenditure, household food use, individual food intake, and program participation--and the 1977-78 NFCS and 1979-80 SFC-LI are now more than a decade old. The table also shows that other databases include some of the types of food consumption data provided by the 1987-88 NFCS:

- ***Household Food Expenditure Data.*** A number of recent databases contain information on aggregate food expenditure.
- ***Household Food Use Data.*** Only the NFCS collects household food use data on a regular basis. The Food Stamp Cashout Evaluations collected information on household food use, but not from nationally representative samples.
- ***Individual Food Intake Data.*** Data on individual food intake are included in a number of large databases, but among recent surveys, only the 1987-88 NFCS and the Continuing Survey of Food Intakes by Individuals (CSFII) contain such data for more than one day. At least two days of intake data are required to obtain unbiased estimates of the proportion of the population that is nutritionally at risk.

The following sections describe databases that could be used as substitutes for the 1987-88 NFCS. For each database, we describe its strengths and weaknesses in meeting FNS' research needs. The Appendix provides abstracts of the 1987-88 NFCS and 10 of the databases that might potentially substitute for the 1987-88 NFCS.

A. THE CONTINUING SURVEY OF FOOD INTAKES BY INDIVIDUALS (CSFII)

The Continuing Survey of Food Intakes by Individuals, sponsored by HNIS, is designed to provide a continuing source of food consumption data in the years between the decennial NFCS. Begun in 1985, the CSFII was fielded for two years before a "hiatus" in 1987 and 1988 when the

TABLE IV.1

CHARACTERISTICS OF DATABASES ON FOOD CONSUMPTION

Database	Sample Size	Disaggregated Food Expenditure Data	Household Food Use Data	Individual Food Intake Data	Participation in the FSP	Participation in WIC	Participation in the NSLP and the SBP
1987-88 NFCS	4,495 hds (basic) 2,509 hds (low-income)	Yes	Yes	3 days	Yes	Yes	Yes
1989 Continuing Survey of Food Intakes by Individuals	1,490 hds* (basic) 725 hds* (low-income)	No	No	3 days	Yes	Yes	Yes
National Health and Nutrition Examination Survey III	35,000 individuals (projected sample size)	No	No	1 day	Yes	Yes	Yes
Consumer Expenditure Survey	20,517 observations (Interview) 11,735 observations (Diary)	Yes	No	No	Yes	No	Yes
1977-78 NFCS and 1979-80 Survey of Food Consumption in Low-Income Households	14,930 hds (basic NFCS) 4,623 hds (low-income NFCS) 3,002 hds (SFC-LI)	Yes	Yes	3 days (NFCS) 1 day (SFC-LI)	Yes	Yes	Yes
Survey of Income and Program Participation	21,206 hds (1990 panel)	No	No	No	Yes	Yes	Yes
Panel Study of Income Dynamics	7,114 hds (1988)	No	No	No	Yes	No	No
Diet and Health Knowledge Survey	1,280 hds* (basic) 626 hds* (low-income) (1989)	No	No	Can be obtained from the CSFII			
Grocery Store Scanner Data	Varies ^b	Yes	No	No	No	No	No
Food Stamp Cashout Demonstration Evaluations	1,000-2,300 hds ^c	Yes	Yes	No	Yes	Yes	Yes
The School Nutrition Dietary Assessment Study	3,500 children, 575 schools (projected sample size)	No	No	1 day	Yes	No	Yes

hds: households

*Preliminary counts

^bOne commercially available scanner data set is based on a sample of 15,000 households; another is based on a sample of 70,000 households^cDepending on the survey

1987-88 NFCS was fielded. It recommenced in 1989 and was fielded for three years until 1991. The CSFII is scheduled to be fielded again between 1993 and 1996.

1. Description of the CSFII

The principal samples for the 1985 and 1986 CSFIIs consisted of women between the ages of 19 and 50 and their children between the ages of 1 and 5 residing in private households in the 48 contiguous states. A secondary sample for the 1985 CSFII consisted of men between the ages of 19 and 50. In 1989, the target population of the CSFII was broadened to coincide with the target population of the 1987-88 NFCS. Hence, the samples for the 1989-91 CSFII include men, women, and children of all ages in the 48 contiguous states.

Each CSFII has included a low-income sample and a basic sample of households. As in the 1987-88 NFCS, the low-income sample contained households whose gross income during the previous month did not exceed 130 percent of poverty, while the basic sample contained households with all levels of income. The principal unit of observation in the CSFII has been the individuals within each sampled household, although each of the surveys has also obtained some household-level information.

The 1989-91 CSFII used a methodology to collect data on individual food intake that is similar to that used in the 1987-88 NFCS. Both surveys collected data on the food intake by each individual in the household on three consecutive days--the day before the interview, the day of the interview, and the day after the interview. Intake data from the first day was collected by recall; data from the second and third days were collected from information recorded by the individuals. In contrast, each respondent in the principal samples for the 1985 and 1986 CSFIIs was asked to recall his or her food intake over the previous 24 hours on six nonconsecutive days at a rate of about one day every two months.¹ The fact that the data were not collected on consecutive days ensured that for each

¹The first day of intake data was collected by in-person interview; the remaining five days were collected by telephone interview if the respondent had a telephone in their home and by in-person interview if they did not. Only one day of intake data were collected from the 1985 secondary sample of men.

respondent, each season of the year was represented. Although the 1989-91 CSFII collected data on consecutive days, it ensures that each season is represented by asking each sampled household and its members to respond at a different time of year.

A major difference between the CSFII and the 1987-88 NFCS is that the CSFII does not collect any information on the price, quantity, or types of food used by the entire household. The only food consumption data collected by the CSFII at the *household* level is the household's usual aggregate expenditure on all types of food over a typical week (or month)² during the previous three months.

The CSFII collects detailed demographic and socioeconomic information, including household income and cash assets, and the age, sex, and race/ethnicity of each individual in the household. It also collects most of the same detailed information about participation in USDA food assistance programs (the Food Stamp, WIC, school lunch and breakfast, and commodity donation programs) as collected by the 1987-88 NFCS. However, the 1989-91 CSFII collects less detailed information on household income than does the 1987-88 NFCS. For instance, the 1987-88 NFCS obtained data on 20 separate sources of household income for each person in the household age 15 years or older. In contrast, the 1989-91 CSFII obtains data on income that are disaggregated by only eight separate sources and that are for the household as a whole rather than for its individual members.

The sizes of the basic and low-income samples of the CSFII are much smaller than the corresponding samples in the 1987-88 NFCS. Final counts of the number of observations in the 1989 CSFII are not yet available, but a preliminary count of the number of the respondents suggests that the basic sample contains about 1,490 households and the low-income sample, about 725 households. These samples are less than one-third the size of the corresponding samples in the 1987-88 NFCS. Data from the 1989 CSFII are scheduled to be available in the fall of 1992.

²The respondent is given the choice of reporting a weekly or monthly figure.

Final response rates for the basic and low-income samples of the 1989 CSFII are not yet available. Preliminary estimates indicate that they will be substantially higher than those obtained for the 1987-88 NFCS.

2. Combining 1989-91 CSFII Data to Obtain More Observations

Data from the CSFII can be combined in either of two ways to obtain more observations on households and individuals: (1) data for the basic and low-income samples for a single year can be combined in one file, and (2) data for the three survey years can be combined in separate basic and low-income files. The latter approach would be easier than the former to implement and would result in a larger combined data file.

a. Combining Data for Different Samples for the Same Year

FNS has twice sponsored the combining of data for the basic and low-income samples of the CSFII. The first combined file contains a single day of food intake data from the 1985 CSFII (Fraker and Post, 1987); the second contains four days of intake data from the same survey (Fraker and Post, 1989). The principal advantage to FNS of sponsoring the construction of a file containing combined data from the basic and low-income samples for a single year of the 1989-91 CSFII would be that the file would provide roughly 60 percent more observations on low-income individuals than are available in the low-income sample alone.³ There would be two principal disadvantages to FNS of following this course. First, FNS' experience is that the construction of this kind of combined file can be time

³Fraker and Post (1989) found that 26 percent of the women and 37 percent of the children in the basic sample of the 1985 CSFII who provided four days of food intake data resided in households with incomes no greater than 130 percent of the poverty guidelines. On the basis of this information, we assume that roughly 30 percent of the 2,925 individuals in the basic sample of the 1989 CSFII who provided three days of intake data were residing in households with incomes no greater than 130 percent of poverty. Thus, we estimate that 878 individuals in the basic sample were in low-income households. In a file containing combined data from the basic and low-income samples of the 1989 CSFII, these 878 cases would increase the number of observations on individuals in low-income households by 64 percent over the 1,372 observations that are available in the low-income sample alone. See the Appendix to this report for further information on sample sizes for the 1989 CSFII.

consuming and expensive.⁴ Second, even with the additional observations on individuals residing in households with incomes no greater than 130 percent of poverty, the combined file would provide data on only approximately 2,250 such individuals, which is less than half of the 5,309 observations provided by the low-income sample of the 1987-88 NFCS.

It appears that it will be unnecessary for FNS to arrange for the combining of data from the basic and low-income samples for a single year of the 1989-91 CSFII, since HNIS plans to provide weights for the combined samples of the 1989 CSFII at both the household level and the individual level. It will be important for FNS to examine HNIS' methodology for computing the combined weights in order to understand any special limitations of estimates based on the combined samples.

b. Combining Data for Comparable Samples for Different Years

By combining data for the low-income samples for the three years of the 1989-91 CSFII, we could generate a single data file that provides three days of food intake data on approximately 4,100 individuals. This is 80 percent more than the 2,250 observations on individuals in low-income households than we would expect in a file containing combined data for the basic and low-income samples of the CSFII for a single year, but it is still substantially fewer observations than are in the low-income sample of the 1987-88 NFCS.⁵

The combining of data for comparable samples for the three years of the 1989-91 CSFII would be much easier and less expensive than the combining of data for the basic and low-income samples for a single year. This is because the data would be from independent samples with the same

⁴The fundamentally different designs of the basic and low-income samples of the 1989-91 CSFII would account for most of the time and expense of combining data from the two samples. The most important design difference is the overrepresentation of households living in high-poverty areas in the low-income sample but not in the basic sample.

⁵The combined basic samples of the 1989-91 CSFII would provide three days of intake data for about 8,800 individuals, whereas the basic sample of the 1987-88 NFCS provides three days of intake data for 8,468 individuals. Our estimates of the combined sample counts for both the basic and the low-income samples of the 1989-91 CSFII are based on the assumption that the sample sizes for the 1990 and 1991 surveys will be approximately equal to the 1989 sample sizes.

fundamental design. We believe that the only required reweighting of the combined data would be to divide the original sample weights by 3. In contrast, the combining of data for the basic and low-income samples would require recomputing the sample weights to adjust for (1) the overrepresentation of low-income cases in the combined data file and (2) the greater variance of estimates based on a sample in which the probabilities of selection differed from cases to case (the low-income sample) as opposed to a sample in which the probabilities of selection were the same from one case to case (the basic sample).

While this approach to combining data from the CSFII would be easy to implement and would provide observations on food intake by a relatively large number of individuals, it would have two disadvantages. First, unlike the first approach, this approach would not produce a file that contains data on individuals from households with incomes no greater than 130 percent of poverty *and* individuals from households with incomes between 130 and 185 percent of poverty. This could complicate analyses of the dietary impacts of programs, such as WIC, that are targeted at individuals residing in households with incomes no greater than 185 percent of poverty. Second, the reference period for the combined data file would be the full three-year period of 1989 through 1991. As a result, it might be difficult to use the combined data to estimate the impacts of program changes that were adopted in the middle of that period. More generally, in using the data, analysts would have to be aware that the long reference period might require them to qualify their findings to reflect this period. Finally, it would be necessary to adjust reported income and expenditure amounts to correct for inflation over the reference period.

3. Strengths and Weaknesses of the CSFII in Meeting FNS' Research Needs

The 1989-91 CSFII has three major strengths in meeting the research needs of FNS:

1. The survey collects food intake data for several days on each individual in the household. Hence, the CSFII data can be used to estimate the distribution of usual nutrient intake and the proportion of the population that is nutritionally at risk.

2. It collects detailed information on participation in a range of USDA food assistance programs.
3. It collects information on a sample of households whose gross monthly income does not exceed 130 percent of poverty. This is the population of households who, because they are income-eligible for food stamps, are of most interest to FNS.

Despite these strengths, the CSFII has two major weaknesses:

1. The survey does not collect information on household food use or on the money value of food used.
2. The sample sizes for each year of the CSFII are small. This limits the capacity of the data to support analyses of subgroups of the population. However, this problem can be addressed to some extent by combining the low-income and basic samples of each CSFII and/or by using more than one year of CSFII data.

B. THE NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY (NHANES)

The National Health and Nutrition Examination Survey is conducted by the National Center for Health Statistics (NCHS) of the Department of Health and Human Services (DHSS). It is designed to collect information on the health characteristics of the U.S. population and the prevalence of behavior affecting the risk of disease. NCHS has completed two national health and nutrition examination surveys--NHANES I, which covered the years 1971 to 1974, and NHANES II, which covered the years 1976 to 1980. A third national survey, NHANES III, began in 1988 and will end in 1994. In addition, NCHS has conducted smaller health and nutrition surveys of special populations, including the Hispanic Health and Nutrition Examination Survey (HHANES) that covered the years 1982 to 1984.

1. Description of NHANES

The target population for the national health and nutrition examination surveys has expanded over time. The sample frame for NHANES I was civilian, noninstitutionalized persons between the ages of 1 and 74 years in the 50 states and the District of Columbia. For NHANES II, the lower age limit was reduced to 6 months. For NHANES III, the lower age limit was reduced to 2 months, and

the upper age limit was eliminated. The HHANES sample frame was persons between the ages of 6 months and 74 years who were (1) Mexican-Americans living in Texas, Colorado, New Mexico, Arizona, or California; (2) Cubans living in Dade County, Florida; or (3) Puerto Ricans living in New York, New Jersey, or Connecticut.

The NHANES surveys are designed so that certain subgroups of the population defined by age, sex, race, and ethnicity are represented by enough persons in the sample to allow subgroup analysis. NHANES III is designed to oversample children, older Americans, African-Americans, and Mexican-Americans. Unlike the NFCS, NHANES does not have a separate low-income sample.

Each respondent is first asked to complete an interview at their home. At this home interview, an appointment is made for the respondent to have a physical examination at a specially designed mobile examination center (MEC). The home interview collects information about the respondent's family as well as about his or her health and risk of disease. At the MEC, the respondent completes a second interview and a series of physical tests (such as physical measurements, blood pressure readings, and dental examinations), biochemical tests (such as urinalysis), and hematological tests.

NHANES III collects two types of food consumption data: frequency of consumption of specific groups of foods and one day of individual food intake. Food frequency data are collected from all respondents age 12 years or older at either the home interview or the MEC interview, depending on the age of the respondent. At the home interview, all respondents age 17 and older are asked how frequently over the past month they consumed food in 60 different groups. An example of one food group is "pork ham, roast pork, pork chops, and spare ribs." Similar food frequency questions are asked of persons between the ages of 12 and 16 at the MEC interview. Also at the MEC interview, parents are asked how frequently they feed their infants (ages 2 to 11 months) certain foods.

The individual food intake data are collected during the MEC interview. All respondents are asked to recall their intake of food over the past 24 hours, including the exact food item, the quantity

consumed, the time and place it was eaten, and the meal at which it was eaten.⁶ Respondents who are 50 years of age or older are contacted by telephone at two later dates and asked to recall their food intake over the past 24 hours. Thus, a total of three days of intake data are collected from respondents ages 50 and older. However, only one day of data on food intake is collected from respondents who are younger than 50.

NHANES III also contains information on demographic and socioeconomic characteristics, health status, utilization of health care, and participation in welfare programs for each respondent. It includes information on total family income (within ranges of \$1,000) during the previous month and previous 12 months, but it does not provide a breakdown of income by individuals in the family. NHANES III also collects some basic information about participation in USDA food assistance programs, including whether anyone in the household currently receives food stamps or has received food stamps during the past 12 months. But unlike the 1987-88 NFCS, it does not include information about the value of the benefits or who is covered by the benefits. Similarly, while school-age respondents are asked whether they attend a school that serves complete breakfasts and lunches and how frequently they eat the meals, they are not asked (as they are in the 1987-88 NFCS) how much they pay for the meals. NHANES III collects information about WIC participation, including whether any person in the household received WIC benefits in the previous month, whether the respondent (age 17 or older) is currently receiving WIC benefits, whether she received WIC benefits during the 12 months preceding the interview, and the length of time she has received WIC benefits.

The samples for the NHANES surveys are much larger than those for the NFCS. NHANES II contains information on more than 20,000 individuals, about twice as many participants as in the basic sample of the 1987-88 NFCS. NHANES III is projected to contain information from interviews on about 35,000 individuals. The response rate is also much higher in the NHANES than in the 1987-88

⁶If the respondent was asked to fast before the examination, the interviewer asks about food intake during the 24 hours before the beginning of the fast.

NFCS. As of April 1991, about 77 percent of eligible individuals completed both the interview and examination components of NHANES III.

One problem with NHANES is that the interviews and examinations do not occur randomly throughout the year. Because of constraints in the operation of the mobile examination center, examinations are performed primarily in the warmer months in states with cold winters and primarily in the cooler months in states with hot summers. Since food intake varies by region and season of the year, this arrangement may lead to biased estimates of average food intake.

To make NHANES III data available as soon as possible, the survey is divided into two phases, each with an independent sample. The first phase covers 1988 to 1991, and the second phase covers 1992 to 1994. Data from the first phase will be available in 1992.

2. Strengths and Weaknesses of NHANES III in Meeting FNS' Research Needs

NHANES III has two major strengths in meeting the research needs of FNS:

1. The survey has a large sample size. NHANES III will include over three times as many individuals as the basic sample of the 1987-88 NFCS. The oversampling of certain population groups will enable researchers to perform subgroup analyses. However, the subgroup of most interest to FNS--individuals from low-income households--is not directly oversampled.
2. It contains one day of food intake data and, although less important to FNS, food frequency data.

NHANES III has three major weaknesses that limit its usefulness:

1. The survey contains no information about household food use or household food expenditures.
2. It contains only one day of food intake data except for individuals 50 years of age and older. Thus, researchers cannot estimate the distribution of usual nutrient intake.
3. Although the survey contains some information about participation in the food stamp, WIC, and school lunch and breakfast programs, it is not sufficiently detailed to support any sophisticated analyses (for example, NHANES III provides no information on the value of a recipient's food stamp benefit).

C. THE CONSUMER EXPENDITURE SURVEY (CES)

The Consumer Expenditure Survey, sponsored by the Bureau of Labor Statistics and conducted by the U.S. Bureau of the Census, is designed to collect data on expenditures by households. One of its purposes is to provide data for updating the Consumer Price Index. The survey was first conducted between 1888 and 1891, and then approximately every ten years afterward until 1980. Since 1980, the survey has been conducted continuously.

1. Description of the CES

The CES has two components: (1) the Interview Survey and (2) the Diary Survey. The Interview Survey collects information about large expenditures, such as automobiles and major appliances, and smaller regular expenditures, such as rent and utility bills. The Diary Survey collects detailed information about frequently purchased small items, such as food, beverages, and personal care products. The two components have different, independent samples and different data collection methodologies.

The target population for both component surveys is the civilian, noninstitutionalized population in the 50 states and the District of Columbia. The Interview Survey has a rotating panel design in which respondents are interviewed every quarter for five quarters. Each quarter, approximately 20 percent of the respondents are removed from the sample and replaced by new respondents. The Diary Survey has a new sample each year, throughout which respondents are interviewed.

The CES samples contain about the same number of households as the actual, as opposed to the planned, number of households in the basic sample of the 1987-88 NFCS. Each year, the Interview and Diary Surveys each contain between 5,000 and 6,000 households.⁷ However, because each

⁷Strictly speaking, the unit of observation in the CES is a consumer unit. A consumer unit is defined as (1) all members of a household who are related by blood, marriage, adoption, or other legal arrangement, (2) persons living alone or sharing a household with others, and (3) two or more persons who live together and pool their income to make joint expenditure decisions. A consumer unit is similar to, but not exactly the same as, a household as defined by the NFCS. The NFCS defines a household as consisting of all individuals who live in the same residence and share a common food supply.

household in the Interview Survey is interviewed four times a year, the interview sample contains about 21,000 observations. The Diary Survey collects data from two consecutive weeks for each household. If each of the two weeks of data is counted as an independent observation, the Diary Survey contains about 12,000 observations. The CES does not contain a low-income sample. In 1990, about 86 percent of eligible sampled households were interviewed in the Interview Survey and about 87 percent of eligible sampled households were interviewed in the Diary Survey. Data from the 1990 CES are now available.

a. The CES Interview Survey

The Interview Survey collects data on food and nonfood expenditures. It provides information about aggregate expenditure on food purchased at the grocery store and food consumed away from home. The respondents are interviewed five times at three-month intervals and are asked to recall their usual expenditure over the previous three months.⁸ It is noteworthy that the respondents report their expenditures aggregated over all foods; the survey provides no data on expenditures on individual food items. The Interview Survey provides information on whether a responding household received food stamps over the previous three months and the value of those benefits, and whether any child in the household purchased school meals during the previous three months and his or her usual weekly expenditure on school meals, but it does not provide information on participation in WIC.⁹

b. The CES Diary Survey

The Diary Survey provides more detail on food expenditures than does the Interview Survey. Two weekly diaries are completed on consecutive weeks. An interviewer visits each household three

⁸In the first of the five interviews, the respondents are asked to recall their expenditure on food over the previous month.

⁹Because the CES is an *expenditure* survey, it provides no information on school meals that were received for free.

times. At the first visit, he or she conducts a short interview with a representative of the household, provides the representative with a food purchase diary, and explains how to complete the diary. About a week later, at the second visit, the interviewer collects and reviews the first diary and leaves a second diary. About a week later at the third visit, the second diary is collected and reviewed.¹⁰

The Diary Survey collects detailed information on the sources of income for each household member (age 14 and older) during the 12 months preceding the interview. It also collects information on participation in the FSP, including whether the household received food stamps in the 12 months preceding the interview, the number of months the household has received food stamps, and the value of the food stamps received in the previous month. The Diary Survey also asks whether any school meals were purchased over the previous 30 days and the weekly expenditures on these meals.¹¹ It contains no information about participation in WIC.

The Diary Survey collects information on food items purchased, the form in which each item was purchased (such as fresh, frozen, bottled, or canned), and the cost of each item. The Diary Survey does distinguish between food purchased for consumption at home and food purchased for consumption away from home. It also distinguishes between food purchased to give as gifts and food purchased for home consumption.

Prior to 1991, the Diary Survey asked respondents to record the *quantity* of each food item purchased. However, accurate estimates of household food use cannot be obtained from these quantity data for two reasons. First, respondents did not use a common measurement unit when recording the quantities of food that they purchased. For example, some respondents recorded purchasing a certain number of "cans" or "bottles" without stating the size of those cans or bottles. Second, the Diary Survey did not collect sufficiently detailed information on the types of food items

¹⁰If the diary has not been entirely completed at the time of the second and third visits, and if the interviewer has contacted the household within eight days of dropping off the diary, then the diary can be completed by recall during the visit.

¹¹Because the CES is an *expenditure* survey, it provides no information on school meals that were received for free.

purchased. For example, while the Diary Survey collected information on the category "cheese," the 1987-88 NFCS collected information on over 80 different types of cheeses. The absence of detail on the types of food used would make it problematic to compute the nutrients provided by that food. The quantity data were not edited and hence were not provided on any data tape released to the public. In 1991 the Diary Survey stopped collecting data on quantities of food purchased.

2. Strengths and Weaknesses of the CES in Meeting FNS' Research Needs

The CES has two major strengths in meeting FNS' research needs:

1. The Diary Survey contains detailed information on food expenditures. The Interview Survey contains information on aggregate food expenditures.
2. Both surveys in the CES contain information on participation in the FSP and limited information on participation in school nutrition programs.

The CES has three major weaknesses that limit its usefulness:

1. The CES contains no information on individual food intake or household food use and currently contains no information on quantities of food purchased.
2. The number of households in the Diary Survey of the CES is not appreciably larger than the number of households in the 1987-88 NFCS. Moreover, low-income households--the households of most interest to FNS--are not overrepresented in the CES samples as they are in the NFCS.
3. The CES contains no information on participation in WIC.

D. THE 1977-78 NFCS AND THE 1979-80 SURVEY OF FOOD CONSUMPTION IN LOW-INCOME HOUSEHOLDS (SFC-LI)

The 1977-78 NFCS and the 1979-80 Survey of Food Consumption in Low-Income Households (SFC-LI) were conducted under contract to HNIS. Both surveys are broadly similar to the 1987-88 NFCS. However, there are some differences in design among the three surveys.¹² We highlight

¹²The earlier surveys also differ from the 1987-88 NFCS in some aspects of data collection. For example, laptop computers were used to record responses in 1987-88 NFCS, but were not used in the earlier surveys.

those differences and assess their implications for the capacity of the survey databases to meet FNS' research needs.

1. Differences in Survey Design

The target population for the basic sample of the 1977-78 NFCS was identical to the target population for the basic sample of the 1987-88 NFCS. However, the criteria for choosing the low-income samples were slightly different. Households were eligible for the 1977-78 NFCS low-income sample if they were either receiving food stamps or other welfare assistance, or if they had income and assets low enough to meet the FSP eligibility requirements. To qualify for the low-income sample of the 1987-88 NFCS, households only needed to meet the FSP gross income eligibility criteria (130 percent of poverty).

The 1977-78 NFCS was conducted between April 1977 and March 1988. It was a much larger survey than the 1987-88 NFCS. For comparison, Table IV.2 presents the sample sizes of the 1977-78 NFCS and the 1987-88 NFCS. The 1977-78 basic sample contains about 15,000 households and 31,000 individuals--about three times the actual size of the 1987-88 NFCS basic sample. The 1977-78 low-income sample contains about 4,700 households and 13,000 individuals compared with about 2,500 households and 6,500 individuals in the 1987-88 NFCS low-income sample.

The 1977-78 NFCS achieved much higher response rates than did the 1987-88 NFCS. Of households in the basic samples of those surveys, 61 percent completed the household interview in 1977-78, compared with 38 percent in 1987-88. Of individuals in the basic samples, 57 percent provided at least one day of food intake data in 1977-78, compared with 31 percent in 1987-88. Response rates are not available for the low-income sample of the 1977-78 NFCS.

The 1979-80 SFC-LI, a follow-up survey to the 1977-78 low-income NFCS, was conducted between November 1979 and March 1980. It was designed to provide information on changes in food use and dietary adequacy in low-income households following the elimination of the food stamp

TABLE IV.2

COMPARISON OF SAMPLE SIZES IN THE 1977-78 NFCS AND 1987-88 NFCS

Unit of Observation	1977-78 NFCS		1987-88 NFCS	
	Basic Sample	Low-Income Sample	Basic Sample	Low-Income Sample
Households	14,930	4,623	4,495	2,509
Individuals				
With at least one day of intake data	30,770	12,847	10,342	6,443 ^a

SOURCE: The sample sizes for the 1977-78 NFCS are reported in Pao, Sykes, and Cypel (1989). The number of households and individuals in the basic sample of the 1987-88 NFCS and the number of households in the low-income sample of the 1987-88 NFCS were counted from the 1987-88 NFCS (unweighted). The number of individuals in the low-income sample of the 1987-88 NFCS is a preliminary count reported in Hubbard et al. (1991).

^aPreliminary count

purchase requirement in January 1979.¹³ As such, it complements the 1977-78 low-income NFCS, which contains information on households before the elimination of the purchase requirement.

The design of the 1979-80 SFC-LI is nearly identical to that of the 1977-78 low-income NFCS. The target population for the 1979-80 SFC-LI was households that would have been eligible for the FSP under the 1977-78 FSP eligibility rules.¹⁴ However, the 1979-80 SFC-LI differs from the 1977-78 low-income NFCS in two respects. First, while the 1979-80 SFC-LI collected seven days of household food use data, it collected only *one* day of individual food intake data from each household member. The 1977-78 NFCS collected *three* consecutive days of intake data from some members of the household.¹⁵ Second, the 1979-80 SFC-LI has a smaller sample than the 1977-78 low-income NFCS. The 1979-80 SFC-LI contains data on approximately 3,000 households, while the 1977-78 low-income NFCS contains data on about 4,600 households.

Sample response rates for the 1979-80 SFC-LI are not available.

2. Strengths and Weaknesses of the 1977-78 NFCS and the 1979-80 SFC-LI in Meeting FNS' Research Needs

Although the 1977-78 NFCS basic and low-income samples are much larger than the respective 1987-88 NFCS samples, and although the 1979-80 SFC-LI sample is similar in size to the 1987-88 NFCS low-income sample, these older databases are unsuitable for most of FNS' current research needs. Because of changes in prices, income, and food assistance programs, it is unlikely that patterns

¹³Prior to 1979, a participating household was required to use 30 percent of its net income to purchase its allotment of food stamp coupons. All households of a given size received the same allotment of coupons, but the amount that they paid for the coupons varied with their net income. The difference between the coupon allotment and the amount paid was referred to as "bonus coupons." Since 1979, participating households have not had to use 30 percent of their net incomes to purchase coupons; rather, they are issued their bonus coupons at no charge.

¹⁴To account for increases in food prices, the income and asset cutoffs used in the 1979-80 SFC-LI were slightly higher than those used in the 1977-78 low-income NFCS.

¹⁵During spring 1977, three days of intake data were collected from all household members. During the rest of the survey months, three days of intake data were collected from all household members 18 years of age or younger and from half of all household members 19 years of age or older.

of food consumption in the late 1970s resemble current patterns. Also, food preferences, the number and kinds of foods that are available, and the nutrient content of foods may have changed dramatically over the past decade.

E. THE SURVEY OF INCOME AND PROGRAM PARTICIPATION (SIPP)

The Survey of Income and Program Participation is a panel survey that has been conducted on an ongoing basis since 1983 by the U.S. Bureau of the Census. Its purpose is to provide information on the income and program participation of individuals and households in the United States. SIPP does not currently collect any food consumption data. The Bureau of the Census has proposed to collect information on aggregate expenditure on groceries and food consumed away from home and whether the household usually has enough to eat on the SIPP "Well-Being" module to be fielded in Wave 6 of the 1991 panel. However, the Bureau of the Census has not yet received approval to field this module from the Office of Management and Budget.

1. Description of SIPP

The target population of SIPP consists of all civilian, noninstitutionalized persons in the United States. Adults (persons 15 years of age or older) in a primary sample of households are interviewed once every four months for a period of up to two and one-half years (typically eight interviews). Each four-month round of interviews is called a "wave." A new panel of adults is introduced annually in February. The 1990 panel will be interviewed during the period from February 1990 through May 1992. Data from the first two waves of interviews with the 1990 panel are now available.

During each wave, all adults in the primary sample and all other adults with whom they reside at the time of interview are asked a fixed set of questions; some information on younger children in the household is gathered from their parents. These core questions yield detailed data on monthly income sources and amounts, household composition, and participation in the labor force and in government transfer programs. Except for the first wave of interviews, the core questions are

supplemented with questions on topics that vary from wave to wave. The topics are extensive and include welfare history, determinants of program eligibility, and asset holdings.

SIPP provides data on a large sample of households. The 1990 panel includes approximately 18,000 households introduced in February 1990 and approximately 3,500 households introduced in February 1989 (the latter is referred to as the "carryover sample" and was introduced to achieve an oversampling of blacks, Hispanics, and households headed by single women in the 1990 panel). The response rate is also high. In the first wave of interviews with the 1990 panel, over 90 percent of eligible households were interviewed, and over 80 percent of eligible households were interviewed in subsequent waves.

SIPP provides detailed information about participation in the FSP, WIC, NSLP, and SBP, including:

- Which household members were authorized to receive food stamps in any of the previous four months, which household members were covered by each authorized person's food stamps, and the value of each authorized recipient's benefit in each of the previous four months
- Which household members were covered by WIC benefits in any of the four previous months, and the months in which each covered person received benefits
- The number of children in the household who receive free or reduced-price school lunches or breakfasts during the current school year and the child's meal-price status (free, reduced-price, or full price) if any school meals were received

2. Strengths and Weaknesses of SIPP in Meeting FNS' Research Needs

It is clear that SIPP could not substitute for the NFCS in any research project that requires food consumption data. However, as SIPP contains a large sample of households and detailed data on program participation and household characteristics, SIPP has been and will continue to be of use to FNS in conducting research on participation in USDA food assistance programs.¹⁶

¹⁶Martini (1992) used SIPP to analyze the determinants of FSP participation among eligible households.

F. THE PANEL STUDY OF INCOME DYNAMICS (PSID)

The Panel Study of Income Dynamics is an ongoing longitudinal survey that provides data on changes over time in labor force participation, employment, income, and the health status of the U.S. population. It is conducted annually by the Institute for Social Research at the University of Michigan under contract to DHHS.

1. Description of the PSID

The PSID was first conducted in 1968. Respondents from 4,802 nationally representative households were interviewed. The survey oversampled low-income households. In 1968, 1,872 households (39 percent) were in the low-income sample, and 2,930 households (61 percent) were in the main sample that included households from all income levels. In subsequent years, interviewers recontacted the households in the 1968 sample and any new households that were formed by individuals who lived in a participating household in 1968. Also contacted were individuals who lived in a participating household in 1968 and later joined other households. This follow-up process has allowed the sample to grow over time even though some households left the sample in later years (because, for example, they no longer wished to participate in the survey). In 1988, respondents from about 7,100 households were interviewed, about half of whom were in the low-income sample. About 76 percent of households contacted in 1968 responded to the 1968 survey. Of households contacted in 1988, 97 percent responded to the survey; of households that were eligible to respond to the 1968 PSID, 37 percent responded to the 1988 survey.

While the PSID collects some information on participation in the FSP, it does not regularly collect information on participation in other USDA food assistance programs. The respondent is asked whether anyone in the household received food stamps in the previous month, and if someone did, the respondent is asked about the value of those benefits. No information is regularly obtained

on participation in the WIC, NSLP, or SBP.¹⁷ The survey collects detailed information about household income during the previous calendar year.

The PSID collects some information on aggregate family expenditures on food for consumption both at and away from home.¹⁸ Households are asked to recall total family *cash* expenditures in an average week on food purchased at a store and consumed at home. In addition, the respondents are asked to recall family expenditures on food delivered to the home and on food consumed away from home.

The PSID does not directly collect data on *total* food expenditures by food stamp recipients. Most analysts using the PSID assume that food stamp recipients spend all their food stamp coupons on food for consumption at home and measure total food expenditures by food stamp recipients by adding the food stamp benefit amount to the reported amount of cash expenditures on food. The advantage of measuring food expenditures in this way rather than from a measure of total food purchases is that it avoids the possibility that the respondent would omit purchases made with food stamps from reported total food purchases. However, it overstates the actual food expenditure to the extent that food stamps are lost, hoarded, or traded for cash or nonfood items. Conversely, it understates actual food expenditure to the extent that food stamps obtained in previous months or from persons outside the household are used to purchase food.

The latest year of available PSID data is 1988.

2. Strengths and Weaknesses of the PSID in Meeting FNS' Research Needs

The PSID may be used to provide data on aggregate household food expenditures. It may be preferable to the CES Interview Survey in some FNS applications because, although the total PSID sample is smaller than the CES sample, the PSID oversamples low-income households.

¹⁷The 1985 PSID asked respondents about participation in the WIC program.

¹⁸During the period 1988-89, the PSID did not collect any information on food expenditures. It resumed the collection of food expenditure data in 1990.

In other respects, the PSID is of limited use to FNS. It does not provide any data on disaggregated household food expenditures, household food use, or individual food intake. Moreover, the food expenditure data may overstate or understate actual food purchases made by food stamp recipients. The absence of routinely collected information on participation in USDA food assistance programs other than the FSP further limits its usefulness.

G. THE DIET AND HEALTH KNOWLEDGE SURVEY (DHKS)

First conducted in 1989, the Diet and Health Knowledge Survey is the first national survey designed to determine how knowledge and attitudes about healthy eating affect food and nutrient intake. Sponsored by HNIS, it is conducted each year as a follow-up survey to the CSFII.¹⁹ Creating a data set that includes both DHKS information on food knowledge and attitudes and CSFII information on individual food intake, household characteristics, and participation in USDA food assistance programs is a straightforward process.

1. Description of the DHKS

All households in the basic and low-income samples of the CSFII that complete an interview are eligible to participate in the DHKS. The main meal planner or preparer from each household is contacted about ten weeks after the completion of the CSFII interview. The interview for the DHKS is conducted by telephone or, if the household does not have a telephone, in person. In 1989, individuals from approximately 1,900 households were interviewed--about 86 percent of all households that participated in the 1989 CSFII.

During a 30-minute interview, the main meal planner or preparer is asked a series of questions about nutrition, attitudes toward food and nutrition, and purchasing and food preparation behavior. The following are some types of information collected by the DHKS:

¹⁹In 1989 only, the Food Safety and Inspection Service of USDA and the Food and Drug Administration of DHHS also sponsored the DHKS along with HNIS.

- Awareness of and attitudes about the dietary guidelines established by USDA and DHHS. For example, respondents are asked how important they believe it is to limit their consumption of cholesterol.
- Methods of food preparation. For example, respondents are asked whether they add salt to food during cooking.
- Use of nutrition information on food labels. For example, respondents are asked how frequently they examine the list of ingredients on packages.
- Knowledge and attitudes about food safety issues. For example, respondents are asked whether they believe it is safe to eat foods that have been treated with irradiation.
- Knowledge of the nutritional content of foods. For example, respondents are asked to compare the fiber content of pairs of foods.
- Knowledge of the relationship between diet and health. For example, respondents are asked about health problems related to excessive fat intake.
- Respondents' perceptions about the healthfulness of their diets.

Data from the 1989 DHKS, combined with data from the 1989 CSFII, will be available in the fall of 1992.

2. Strengths and Weaknesses of the DHKS in Meeting FNS' Research Needs

There are two major strengths of the DHKS. First, it can provide FNS with information on knowledge and attitudes about nutrition that is not currently available from any other database.

Second, when linked with the CSFII, it will enable FNS to examine such issues as:

- The relationship between knowledge and attitudes about nutrition and actual nutrient intake
- The effect of nutrition education programs on knowledge and attitudes about nutrition and on the intake of food and nutrients
- The relationship between knowledge and attitudes about nutrition and participation in food assistance programs

The major weakness of the DHKS is that it contains no information on household food use. While data on individual food intake can be added to the database by linking it with the CSFII, the CSFII does not contain data on household food use.

H. GROCERY STORE SCAN DATA²⁰

A large proportion of supermarkets, perhaps as high as 90 percent, use scanner systems in the checkout process.²¹ Central to these systems are optical scanners which translate the bar codes on product packages into data that can be processed by computers. The detailed nature of these data permit the identification of the type of product, brand, size, and price of each item being purchased. These systems transmit this information to cash registers, thereby eliminating the need to enter the information manually. This reduces checkout time and errors and provides customers with more detailed receipts. While scanner systems improve both productivity and efficiency in the checkout process, the data that they generate can, if appropriately stored and analyzed, be a valuable source of information on consumer shopping behavior.

This section describes scanner systems, types of scan data that are available, and caveats associated with the various levels of record keeping.²² These descriptions are based on telephone conversations between the authors of this report and a number of individuals who are knowledgeable about scanner systems and databases related to those systems and on written input from Dr. Eastwood. Because no single individual with whom we spoke had comprehensive knowledge of this topic, we had to synthesize the information we obtained. As a consequence of this process, our

²⁰Dr. David Eastwood, Professor of Agricultural Economics and Rural Sociology at the University of Tennessee, is a co-author of this section. Dr. Eastwood is the director of the University of Tennessee's Scan Data Project, which focuses on the development and analysis of databases derived from retail scanner checkout systems.

²¹While roughly 90 percent of supermarkets have scanner systems, substantially smaller percentages of other types of food stores have them.

²²We use the term "scanner" to refer to refer to integrated retail checkout systems and to their operation. We use the term "scan" to refer to the output of such systems. Thus, we refer to "scan data" being obtained from "scanner systems."

information on scanner systems and databases is less precise, less complete, and less authoritative than is the information on the other databases described in this chapter.

We spoke with individuals employed by the following types of organizations:

- A major supermarket chain (one individual)
- Two policy research firms with expertise in electronic benefit transfer and scanner data systems (three individuals)
- A federal government research institute (one individual)
- A FNS regional office (one Individual)
- A food marketing trade association (one individual)
- Two market research firms that provide scanner databases to clients (multiple individuals in each firm)

In addition, this section was reviewed by Dr. Eastwood, who also contributed to some of the subsections.

1. General Information on Scanner Systems

The Universal Product Code (UPC) has become the standard, both nationally and internationally, for identifying manufacturers and their products. The objective is to have a numbering system that allows for the identification and tracking of individual goods, including foods, from their production by specific manufacturers through sales to consumers.

a. The UPC Format for Retail Products²³

Guidelines for UPCs are controlled by the Uniform Code Council, which sets the rules for several types of product numbering systems. Since FNS's interests center on food consumption, the retail-level numbering system is outlined below. UPCs have two visual components. One is a set of arithmetic numbers, which people can read. The other is a set of vertical bars that vary in width and

²³This section is based on: Uniform Code Council, Inc. (1991).

spacing, which can be "read" by optical scanners. By convention, the arithmetic numbers appear at the bottom of the UPC label. The pattern of bars and spaces interpreted by the optical scanners yields the same total numerical value as the set of arithmetic numbers. However, the former is an electronic record that can be processed by computer systems. It is the automatic conversion to electronic records that makes the UPC format so powerful in today's computer-oriented world.

While the concept is straightforward, its implementation requires universal agreement on the part of manufacturers and retailers on the form of the numbering system. With respect to the food sector, there are some special problems that must be accommodated. One is that some foods are not processed to a significant extent (e.g., fresh produce). Consequently, there are many growers and other economic agents involved in the distribution and sale of such products, and identifying each may not be warranted, let alone feasible. Another problem is that the retail sizes of some foods are variable (e.g., fresh meat), so the UPC must be flexible enough to allow for an indefinite number of different sizes without denoting different products. A third complicating factor is that grocery stores prepare items in each outlet and sell them directly to customers. This situation creates a need for being able to keep some codes under the purview of individual outlets.

Reading from left to right, the first six digits in a twelve-digit code are assigned by the UPC Council. The extreme left-hand digit denotes the product category. Those that pertain to food are 0, 6, 7 for regular codes; 2 for variable-weight items; 4 for instore labeling; and 5 for coupons. Except in the case of variable-weight items (discussed below), digits 2 through 6 identify the manufacturer. The next five digits are codes assigned by the individual manufacturers to identify specific products. The final digit is a number that is based upon the preceding 11 according to a formula. Scanner systems use this digit to verify that they have properly scanned and interpreted the UPC.

The manufacturer-assigned codes (digits 7 through 11), while unique numbers for individual food items, are not assigned in a way that permits sorting according to food groups. Also, one cannot group UPCs by store department (produce, frozen foods, etc.) because supermarkets do not group

foods in identical ways. The lack of comparability is the result of the historical development of each outlet and of management-labor negotiations. This means that any organization of data based on UPCs requires the generation of a master list of foods and their associated codes. Consequently, considerable effort is required at the retail level to organize scan data into meaningful food groups from the consumer's perspective and for food policy analyses.

The sizes of fixed-weight products are not addressed in a unique part of the UPC. Rather, manufacturers simply assign different numbers in the five-digit product fields to denote both various products and various sizes of the same product. Therefore, the same processed food sold in different sizes is assigned different numbers by its processor according to the processor's internal numbering scheme.

Given the flexibility in the numbering system and the nature of the special codes described below, it is not possible to determine the number of food codes being used at the present time at the retail level.

Once an optical scanner has read a bar code during the checkout process, the code is matched with an identical code in a price file. The latter contains the product's unit price, description, size, and (depending on a store's computer software) department. This information is displayed visually at the register, appears on the customer's receipt, and becomes part of the sales record that can be stored for subsequent processing. If the customer has informed the checkout clerk that payment will be made in full or in part with food stamps, the clerk can request that the cash register automatically ring up one total purchase amount for food stamp-eligible items and one for ineligible items. Cents-off coupons can also be read by the scanners and incorporated into the calculation of the total purchase amount.

Note that product descriptions are not a formal part of the UPC format. Rather, they are developed and maintained by individual retailers in their price files, as just described. A retailer uses

the UPC to link to its price file and to retrieve the name or description of a product as it is scanned during the checkout process.

b. Manufacturers of Scanner Equipment

Many manufacturers produce individual components of scanner systems; but only about half a dozen produce complete integrated systems, consisting of an optical scanner, scale, checkout terminal, and associated software and data files. NCR and IBM are two of the largest manufacturers of integrated systems. Currently, the retail food checkout component of the market for integrated systems is experiencing some new entrants.

Incompatibility of scanner systems across different outlets of a multi-outlet store or with other computer systems maintained by a store can be a problem. There is no assurance that scanner hardware and software are compatible across manufacturers or are compatible with other computer systems used at the corporate management level. The only assured consistency across scanner systems is the UPC. For example, a representative of a major supermarket chain in the Southwestern United States informed MPR that the company uses scanner equipment from five different manufacturers. Another example is a chain in the Southeast that had a scanner system that was incompatible with management's computer system, so there was no way to retrieve the scan data from the computers operating the scanners and cash registers. It therefore was not possible to use the scan data to inform management decision making.

c. Variable-Weight Foods

Variable-weight foods are either sold without being packaged (e.g., grapes) or are packaged in such a way that the weights of individual packages are not fixed (e.g., rib roast). These foods are concentrated in the 1) fresh produce, 2) fresh meat and fish, 3) service deli, and 4) instore bakery

departments of supermarkets. In 1990, food items in these four departments accounted for 39 percent of total retail food sales (*Supermarket Business*, September 1991, page 61).¹⁸

The initial UPC digit for variable-weight foods is 2, and the format of the remaining digits has been specified by the UPC Council. The second digit is a code to denote the type of packer, but it is left to the retailer to assign numbers in some meaningful way. UPC digits 3 through 6 represent values assigned by various trade organizations. Ranges of values of these digits have been established for these organizations by the UPC Council. Each organization, however, may assign numbers within its purview to specific foods. For example, produce codes range from 4011 to 4959 and are assigned by the Produce Electronic Identification Board. The seventh digit is called the price-check digit, which is followed by four digits that are the price of the package. The last digit continues to be a check for scanner accuracy, as is the case for all UPCs.

Variable-weight items pose two special problems with respect to generating a scan data base. First, the numbering system is not unique. Packer codes (digit 2) differ across outlets, and trade organizations are in various stages of creating unique item identifications (digits 3 through 6). Consequently, the UPC for a variable-weight food can differ by outlet. Second, the four digits representing the dollar value of a variable-weight food item (digits 8 through 11) pose special software problems for collecting scan data beyond a customer's total dollar purchase amount over all items. Some scanner software is capable of recognizing digits 8 through 11 as representing a dollar value and associating it with the product identifier (digits 3 through 6). Less sophisticated software is not capable of this.

Two types of quantity data exist, depending upon the management information software that is used to process variable-weight data. Software that recognizes digits 8 through 11 as representing the dollar value of a unique item identified by digits 3 through 6 can use price-per-pound information

¹⁸The percentage cited is the percentage of all retail food sales, without regard for the type of store. Excluded are sales of food by restaurants.

in the outlet's price file to calculate the number of pounds of the item that were purchased. The other type of quantity data is called item movement. The management information software in these instances does not use the dollar value information imbedded in digits 8 through 11, along with the price-per-pound data, to calculate the pounds purchased. Rather, it simply keeps a running total of the number of times each unique variable-weight item identifier (digits 3 through 6) is scanned. This software provides information on the number of units of the item that were purchased, but not on the weight of each unit or on their aggregate weight.

d. Private Label Foods

A similar data processing problem arises with UPCs that begin with 0 or 4. The remaining digits are unique to individual outlets even though they are fixed-weight items. Private label foods are examples, and within a typical supermarket they may be found in several departments, such as canned foods.

2. Forms of Scan Data

Scan data are currently analyzed in five different forms. First, optical scanners and checkout scales record customer purchases by product and prepare a receipt. This information is transmitted from the checkout terminals to computer storage devices for subsequent processing by a management information system. Second, many supermarkets also have instore computers that can track some of the store's products, so point-of-sale data in these instances are transmitted to the instore tracking system. At regular intervals the stored point-of-sale data are aggregated across customers, time, and products and are transmitted to corporate headquarters. These aggregated corporate-level data are the third type of scan data. The corporate-level data may be sold to companies that specialize in market research, which then aggregate the data across supermarket chains and independent stores, resulting in the fourth type of scan data.

The fifth form in which scan data are available not only contains information on food quantities and expenditures, but also information on food shopper characteristics generated by smart card applications, frequent shopper programs, or by market research firms. The common features of these sources of data on shoppers are 1) food shoppers complete a questionnaire that gathers socioeconomic data on themselves and their households and 2) they are issued a card that contains an ID that can be matched to the questionnaire. One example of the use of this technology is the enrollment by market research firms of food shoppers in test areas for the purpose of conducting controlled marketing experiments. Participants in such an experiment are instructed to use retail outlets that have agreed to supply their sales data to the coordinating market research firm.

The most sophisticated variant of the fifth type of scan data entails issuing a "smart card" to participating shoppers. This card has an encodable strip that can be read or written on by an electronic card reader that has been integrated into the scanner system. At the end of each transaction, the scan data are written onto the smart card. At the beginning of the first transaction of a new reference period (e.g., a week), the accumulated information on a smart card is read by the card reader and transmitted to that customer's file in a central computer.

Of the five types of scan data, the fifth type has the greatest potential to support research of relevance to FNS. The other four do not provide information about the characteristics of the food shopper. In addition, variations in the management information systems that are used to process the data make cross-chain comparisons tenuous.

3. Uses of Scan Data by Food Retailers

Wide variation exists in the extent to which the data generated by scanner systems are used by food retailers. Small local and more progressive regional chains of supermarkets tend to be the most innovative in their uses of scan data. This section discusses three applications of scan data by food retailers. It concludes by discussing the relevance to FNS of these applications of scan data.

a. Scan Data Used in the Checkout Process

At the lower end of the range of applications of scan data are their use by food retailers to facilitate the checkout process. Many food retailers maintain inventory and accounting systems that are independent of their scanner checkout systems. Their only significant use of scan data is in the checkout process. Dr. Eastwood and our contact at a major Southwestern supermarket report that many individual grocery stores and chains have not progressed beyond this most basic application. In some of these organizations, management has simply have been too busy dealing with day-to-day operations to consider longer-run applications of scan data. In other cases, management has chosen to invest fewer resources in developing the capacity to use scan data in inventory, accounting, and merchandising applications.

b. Scan Data Used in Inventory Control and Related Applications

Relatively straightforward uses of scan data to control inventories and to facilitate other management activities require that the data generated by all sales transactions be continuously aggregated by product (resulting in the second type of scan database discussed above--scan data at the instore-computer-level). The management-information software that has been developed for this purpose is based upon fixed-weight items. In these instances, item movement can serve as an appropriate measure for restocking, and the relatively long shelf-lives and delivery times of fixed-weight foods make them well-suited for inventory management and just-in-time delivery systems. Other applications based on the instore-computer-level scan data include processing transactions by time of day and day of the week to improve labor schedules.

c. Scan Data Used in Marketing and Demand Analyses

Theoretically, scan data permit the tracking of the sales of individual products and the analysis of the sales data by time, store, region, and the characteristics of the purchasers. Marketing and applied demand analysts have begun to tap this potential of scan data. The types of analyses that

have been completed have depended, in part, on the types of scan data available for the research. These have included analyses of the effects of price changes, packaging, newspaper and electronic media advertising, instore merchandising, and generic promotions. Estimates of the effects of changes in the price of a product on sales of that product and on sales of substitute and complimentary products have been generated, along with estimates of the effects of various merchandising campaigns.

d. Relevance to FNS of These Uses of Scan Data

Most current applications of scan data by grocery stores entail the aggregation of the data over customers and transactions by product. Data in this form (unit sales per product per time period) are of little use to FNS in addressing its needs for individual- and household-level data on food consumption. The relatively recent evolution of the micro-level data via frequent shopper and smart card programs point to the potential for generating the requisite data. However, several years of commercial development appear to be required before these data can be collected on a sufficiently broad scale in terms of the number of food shoppers and the types of food outlets to be of value to FNS.

Several key problems will have to be addressed in the evolution of scanner systems and scan databases. First, industry groups that control the designation of codes for variable-weight foods will have to introduce standard codes that would apply across stores. Second, software that uses pounds sold instead of item movement for variable-weight foods needs to become commonplace. Third, all retail outlets for food to be consumed at home must have scan data systems. Fourth, the UPC master files used by the participating outlets must be continuously updated to incorporate new foods as well as to adjust for the effects of mergers and acquisitions on UPC designations.

As these problems are solved, the feasibility of using scan data to address FNS' research needs will increase. Current data collection methods (e.g., for the Nationwide Food Consumption Survey) are very costly and time intensive. Scan data have the potential to render those methods obsolete.

Once smart card applications have addressed the concerns noted above, then consumer purchases of food identified by UPC could be matched with a UPC-specific nutrient data file, in a process similar to that which is currently used to convert food-use data from the Nationwide Food Consumption Survey into data on nutrient availability. Then, a longitudinal data base could be developed to serve as the standard resource for food demand and nutrition-related research of relevance to FNS.

4. Proprietary Data Sets Derived from Scanner Systems

Given the heterogeneity in scanner system hardware and software and problems with cross-store variations in UPCs, FNS would have to make a major investment in order to collect scan data from a nationally representative sample of households and/or grocery stores. An alternative would be to negotiate access to one or more of the proprietary databases that are generated by market research firms. This section briefly describes five such databases--two from the A. C. Nielsen Company, two from Information Resources, Inc. (IRI), and one that is based on frequent shopper programs that have been put in place by a number of supermarket chains and individual stores.¹⁹ It also discusses some features of these databases that may limit their usefulness to FNS.

a. A.C. Nielsen Data Sets

Scantrack. This proprietary data set provides continuously updated monthly and annual sales data (in both item movement and dollar volume) for all fixed-weight food items. Limited coverage of variable-weight foods is provided. The data are obtained from a nationally representative sample of over 2,000 grocery stores and are weighted to represent the total national grocery store sales volume. All data are reported in aggregate form by type of food. Separate data on food items purchased with food stamps are not available.

¹⁹The sources that we contacted at A.C. Nielsen and IRI were unwilling to quote prices of their data products or software technology without much more detailed information about data requirements and plans for using the data.

Household Panel. Beginning in 1989, A.C. Nielsen provided a nationally representative sample of 15,000 households with scanning equipment to use in their homes to record all UPC purchases. These households are located in neighborhoods served by the grocery stores in the Scantrack sample. The household panel is periodically updated, and the participating households annually provide Nielsen with data on the socioeconomic characteristics. Each time a household in the panel scans its food purchases, it keys in the name of the store from which the food was purchased. If that store is a Scantrack store, Nielsen extracts the price information from its Scantrack system files. If it is not a Scantrack store, the household must key in the price information. The household uses the scanning equipment to transmit the food purchase data to Nielsen.

No information is collected on whether the participating households receive food stamps or on whether they use food stamps to purchase specific food items. FNS might be able to negotiate the collection of this information.

b. Information Resources, Inc. Data Sets

IRI has two proprietary scanner-based data sets that are similar to the two Nielsen data sets. That is, one pertains to grocery store sales, and the other pertains to household food purchases.

InfoScan. Each week IRI obtains sales data (in both item movement and dollar volume) on all fixed-weight UPC food items for a nationally representative sample of 3,000 grocery stores. These stores were selected to provide a clustered nationwide representative sample of food and other UPC commodities sold by major supermarkets. Sampling weights are assigned to each store, and weighted aggregates across the sampled stores are generated for each UPC. IRI contends that the resultant totals are valid estimates of grocery store sales nationwide. Socioeconomic characteristics for each store's location are also available. Subscribers are provided with weekly electronic copies of the aggregated data on unit and dollar sales for the total sample and for the subscriber's local cluster and store.

BehaviorScan. IRI has given identification cards to 70,000 households in neighborhoods served by a subset of the 3,000 clustered grocery stores that comprise the InfoScan database. Prior to receiving their cards, food shoppers must complete questionnaires that gather socioeconomic information. All major supermarkets in the clustered areas use scanners and participate in InfoScan. A random sampling process is used to select households in each of the clusters. There is some limited capability to track variable-weight items. The cards are to be presented each time food purchases are made, so the socioeconomic data can be linked with the purchase data.

IRI does not currently determine whether a panel household is a food stamp recipient or whether specific purchases were made with food stamps. However, FNS might be able to negotiate the collection of this information.

c. Frequent Shopper Programs

Some supermarkets have implemented their own shopper card programs that are designed to promote customer loyalty and to provide management with more specific socioeconomic information that can be used in micromarketing applications. Conceptually, this situation is comparable to that of IRI and Nielsen household panels. Some market research firms have recently started to offer smart card services to supermarkets. These firms provide the hardware and software necessary to implement the frequent shopper programs as well as to collect the socioeconomic data.

d. Limitations of Proprietary Data Sets

While the proprietary household-level scan data sets (Nielsen's household panel, IRI's BehaviorScan, and data produced by frequent shopper programs) have several desirable features, there are some key problems that must be solved for them to be suitable to address FNS' research needs. These problems are outlined below. Presently, they preclude the use of the data, but as scan data technology and UPC implementation evolve, FNS may find it useful to draw upon the

experiences and expertise associated with these proprietary data and the researchers who analyze them.

Some of the limitations of the proprietary scan data sets are common to all UPC-based data. Although the codes are very well structured, there are problems when aggregating across stores. Part of the difficulty associated variable-weight foods will go away as industry groups develop, implement, and refine their coding systems. But part will remain because of the need to have the flexibility to generate instore UPCs. Another general problem is that not all food outlets have scanner capability. If complete records of food purchases by individuals or households are to be obtained, then it will be necessary to include purchases made at convenience stores, farmers' markets, and small grocery stores, in addition to purchases made at supermarkets.

Other problems are specific to proprietary data. At present, the vendors of proprietary scan data sets are very reluctant to document their statistical procedures. Among the consequently unaddressable concerns about these data are the usual issues of sampling design, nonreporting, and creation of weights. There also are methodological concerns that are unique to scan data, such as the following:

- Some scan data systems record only item movement while others record pounds sold.
- If a scanner cannot read a UPC, the clerk may have the option of keying in a product category code for the product along with the price. Some stores give much more leeway than others for this to occur. This practice makes it impossible to precisely identify a product.
- As products approach their sell dates, some stores lower the prices and give the products default category codes (e.g., meat), thereby making it impossible to precisely identify the product.

Two principle problems need to be addressed if FNS is to use proprietary scan data sets. The first is a problem of selection bias, which arises because participation on the part of food shoppers in the Nielsen and IRI household panels is voluntary, as is participation in frequent shopper

programs. Consequently, lower-income households and minority groups may be underrepresented in proprietary scan data sets. The situation is complicated further by the greater use of frequent shopper programs by more progressive chains, which tend to cater to more affluent food shoppers. Second, the lack of uniformity in the collection of socioeconomic data, and the absence of unique household identification system, effectively preclude the merging of data from the frequent shopper programs of different food retailers.

5. Food-Purchase Data in an EBT Environment

In food stamp electronic benefit transfer (EBT) demonstration areas, all stores authorized to accept food stamps are provided with equipment that:

- Reads a food stamp recipient's identification number from an EBT card
- Allows the recipient to key in a second number, a "personal identification number," which is known only by the recipient, thus reducing the likelihood that the card could be used by an unauthorized person
- For an on-line system, transmits the food stamp purchase amount to a centralized computer, which debits the cardholder's EBT account and credits the store's account; for an off-line system, debits the balance that is encoded in the recipient's smart card and records a credit for the store.

A store within an EBT demonstration area typically has the option of obtaining its EBT equipment from the demonstration's sponsoring agency or from a third-party vendor.

If EBT systems were linked to scanner systems, and if scan data files included details on individual food items purchased by each customer, then it would be technically possible for researchers to use the linked data to analyze the food purchasing behavior of food stamp recipients as a function of their household characteristics. The latter could be obtained from food stamp program files. However, the two systems are not typically linked, and the scanner systems used by most stores do not include data files that contain information on the food items purchased by individual customers.

It is not necessary to link an EBT system to a store's scanner system; no such link is included in the EBT system basic components described above. Thus, while it can reasonably be expected that a food retailer and the EBT-sponsoring agency will maintain files in which EBT total purchase amounts are stored along with the purchaser's EBT identification number, they will have little incentive to either store in these files data on the food items purchased by EBT participants or to link the EBT files to scan data files that hold the food-item information.

In principle, the technology associated with frequent-shopper programs could be adapted to meet FNS' needs for food-purchase data in an EBT environment. The most likely avenue would entail the use of smart cards that have enough storage capacity to keep a running balance of a participant's food stamp account and to store data on individual food items purchased by UPC. The food-purchase data could be transferred from the cards on a systematic basis (this might occur at the local food stamp office each month at the same time that the cards are credited with the next month's benefit allotment) to construct an FNS food-purchase database that would be comparable to the IRI and Nielsen household-level panels. This approach would avoid the problem of food shoppers patronizing a variety of outlets and having to somehow aggregate across those outlets to obtain each person's food purchases. A major complication of this approach would be inconsistency across stores in codes for variable-weight foods and private label foods.

6. Strengths and Weaknesses of Scan Data in Meeting FNS' Research Needs

Grocery store scanner systems are a potential source of timely data on household food purchasing behavior. Four desirable features of the technology are as follows:

1. The ability to record automatically purchases of individual food products
2. Respondent burden is low
3. Technologies are emerging that will give time and cost advantages to this capability

4. The proliferation of scanners and the evolution of UPCs and smart cards will continue to increase the potential completeness of coverage of scan data files--most notably with respect to variable-weight foods and foods purchased in smaller stores

These positive features of scanner systems may mean that, at some future time, it will be feasible for the USDA to collect comprehensive scan data on food purchases directly or from a vendor. These would be household-level data on individual food items, with socioeconomic data appended. The later could be obtained either from Food Stamp Program records or from a survey of participants in the scan data collection project. Because that survey would gather data only on socioeconomic characteristics, it would be far less burdensome than the NFCS.

However, five features of scanner systems severely limit present applications of the scanner technology by FNS to carry out its research agenda.

1. Scan data bases do not include any information on nonpurchased foods that may be used by a household or on food intake by individual members of a household.
2. Not all retail level food stores have scanner equipment. Approximately 90 percent of supermarkets and smaller percentages of other types of grocery stores have the equipment. Food stores that serve large numbers of food stamp recipients may be less likely to have scanner systems. If so, purchases made by food stamp recipients would be underrepresented in national databases compiled from scanner systems, unless the data were weighted to correct this problem.
3. Not all UPC food codes are unique across stores; these include the codes for variable-weight items, instore-prepared items, and private label items.
4. Considerable programming resources may be required to link scan data on household food purchases with data on household characteristics. Frequent shopper, smart card, and some proprietary databases have this currently have this capability via the use of identification cards; however, this capability does not generally extend beyond one outlet or one supermarket chain.
5. Assignment of UPC-level nutritional values would require extensive modification of existing nutrient databases.

I. DATABASES FROM THE FOOD STAMP CASHOUT DEMONSTRATION EVALUATIONS

A number of different state and local governments have recently conducted or are currently conducting demonstration programs that issue food stamp benefits in the form of checks (cashout)

rather than the traditional coupons. In Washington State, Alabama, and San Diego County, FNS is sponsoring evaluations of the effects of cashout on the use of food at home and related consumption behavior by food stamp households. Central to these evaluations is the collection of seven days of data on household food use from randomly selected samples of coupon and check recipients. While these data sets are designed to meet the research needs of the food stamp cashout evaluations, it is unlikely that they will be useful as principal or supplemental data sources in other, noncashout applications.

1. Description of the Food Stamp Cashout Demonstration Evaluations

FNS sponsored the collection of household food use and related data from 1,184 households in 10 counties in Washington State from August through October 1990; from 2,291 households in 12 counties in Alabama from August through November 1990; and from 1,078 households in San Diego County from May through August 1990. Data were collected from roughly equal numbers of check and coupon households. In all of the demonstration sites in Alabama and San Diego County, subsets of ongoing and new participants in the FSP were randomly selected into the evaluations as check (treatment) cases, while others were randomly selected as coupon (control) cases. In Washington State, entire caseloads within counties were assigned to coupon or check status, and subsets of these caseloads were randomly selected into the sample for the evaluation survey.

Except for relying on paper instruments instead of laptop computers to administer the surveys, the Cashout surveys used essentially the same methodology as the 1987-88 NFCS to collect detailed data on household food use and to convert these data into information on the money value of food used at home and the nutrients available from that food. The Cashout surveys also obtained data on household demographics, income, usual monthly expenditures on food used at home and food used away from home, expenditures on a wide range of nonfood consumer goods and services, attitudes toward the FSP, and eligibility for and participation in USDA food assistance programs other than the FSP. The Cashout surveys did not obtain information on individual food intake.

Public use files containing data from the Cashout surveys should be available in the spring of 1992 for the Alabama and San Diego County evaluations and in the summer of 1992 for the Washington State evaluation.

2. Strengths and Weaknesses of the Food Stamp Cashout Demonstration Evaluation Databases in Meeting FNS' Research Needs

The Cashout evaluation databases were designed to provide an empirical basis for evaluating the effects of check-based food stamp benefits on household food consumption and related behavior. The strengths and weaknesses of these databases flow directly from the fact that they were designed to meet this very specific research need.

The Cashout evaluation databases have six strengths in meeting FNS' *general* research needs for data on food consumption:

1. They provide information on household food use and nutrient availability at the same level of detail as the NFCS.
2. They also provide information on household demographics, income, and participation in USDA food assistance programs at the same level of detail as the NFCS.
3. They provide large amounts of information on nonfood consumption--more than is available in the NFCS, but less than in the CES.
4. Sampled households participated in the Cashout evaluation surveys at rates of 75 to 78 percent, roughly twice the household response rate for the 1987-88 NFCS.
5. They provide food use data on large numbers of households: 1,078 in San Diego County, 2,291 in Alabama, and 1,184 in Washington State.
6. They provide food use data on recipients of both food stamp coupons and food stamp checks; therefore, they are capable of supporting analyses of the effects of checks versus coupons on food use.

The Cashout evaluation databases have four principal weaknesses that limit their usefulness in meeting FNS' *general* research needs for data on food consumption:

1. The samples from the Cashout evaluations are representative not of the national food stamp population, but of the food stamp population in three very limited geographic areas.
2. The evaluations provide no information on eligible nonparticipants in the FSP, thus greatly circumscribing their capacity to support analyses of the effects of food stamps on consumption behavior.
3. They provide no information on individual food intake.
4. Half of the sample observations in the Cashout evaluations--the observations on recipients of food stamp checks--would be of little use in analyses not related to cashout. The remaining sample observations may be too few in number to support conventional subgroup analyses.

J. THE SCHOOL NUTRITION DIETARY ASSESSMENT (SNDA) STUDY DATABASE

As part of its ongoing research agenda, FNS is currently funding or planning to fund research projects that will collect data on food consumption and program participation. Some of the data that will be collected as part of these studies may be useful substitutes for the 1987-88 NFCS. One large-scale research projects is particularly noteworthy--the School Nutrition Dietary Assessment Study.

1. Description of the SNDA Study

The School Nutrition Dietary Assessment Study was designed to examine the impact of USDA school nutrition programs on the dietary status of school children. The study addresses four issues: (1) the nutrient content of USDA school meals and the types of non-USDA food *offered* by schools, (2) the nutrient content of USDA and non-USDA school meals *selected* and *consumed* by children, (3) the impact of the school nutrition programs on the nutrient intake of school children, and (4) whether the impact of the school nutrition programs on the nutrient intake of children has changed since the early 1980s. A similar study funded by FNS, the National Evaluation of the School Nutrition Programs (NESNP-I), collected data on the dietary status of school children during the 1980-81 school year. However, changes in the operation of the school nutrition programs since the early 1980s require the updating of information on their effects.

The target sample for the SNDA consists of 350 nationally representative public and private schools in the 48 contiguous states and the District of Columbia, and 3,500 children in grades 1 through 12 who attend those schools. School-specific data only are being collected from an additional 225 schools. Eighty-five percent of sampled schools are expected to participate in the study and, within those schools, 85 percent of sampled students are expected to participate. Thus, the expected student response rate is 72 percent. The data collection period is January to May 1992.

The SNDA study is collecting seven types of data:

1. ***The Content of USDA Meals Offered by Schools.*** Detailed information on the content of the USDA meals offered during a five-day period are being collected from 600 schools. The information includes menus, recipes, and the nutrient composition of purchased food products. Data are also being collected via a mail questionnaire to school food service personnel and either an in-person or a telephone follow-up interview.
2. ***The Types of Non-USDA Food Offered by Schools.*** This includes food sold at school snack bars and vending machines. The types of non-USDA food available to children during a five-day period is described by research staff visiting 375 schools.
3. ***The Foods Selected and Consumed by Children at School.*** Children from all grades are asked to describe the foods they selected and consumed at the last meal they ate at school. An attempt is made to interview half of the children after breakfast and half of the children after lunch.
4. ***The Dietary Intake of Children.*** Children from all grades are asked to recall all foods they ate in the previous 24 hours. Children in grades 3 through 12 are interviewed at school. Children in grades 1 and 2 are interviewed with their parents in their home.
5. ***The Demographic Characteristics and Family Income of the School Children.*** During interviews, children are asked to provide demographic information including their age, sex, grade, and ethnicity, and the size of their family. Information on family income, and hence eligibility for the school lunch and breakfast programs, is obtained from the children's parents via a mail questionnaire (with a telephone follow-up).
6. ***The Characteristics of the Schools and School Meal Services.*** Descriptive information about the schools and school meal services, including meal prices, are collected (via a mail questionnaire) from school officials and school food service personnel.
7. ***The Children's Eligibility for the School Lunch and Breakfast Programs.*** A list of children in the sample who are certified as eligible to receive free or reduced-price school lunches and breakfasts is obtained from school administrative staff. (Some

children may be eligible to participate in these programs, but are not certified as eligible because their parents did not apply for free or reduced-price school meals.)

2. Strengths and Weaknesses of the SNDA Database in Meeting FNS' Research Needs

For research projects on school nutrition programs, the database from the SNDA Study may be a good substitute for the 1987-88 NFCS. Both collect information on individual food intake, participation in the school nutrition programs, and eligibility for the school nutrition programs. However, while the 1987-88 NFCS collects three days of individual food intake data, the SNDA database will collect only one day of food intake data. Hence, the SNDA database cannot be used to estimate the distribution of usual nutrient intake. On the other hand, as a data source for research on school nutrition programs, the SNDA database has two advantages over the 1987-88 NFCS. First, the SNDA database will contain data on about 3,500 children, while the basic sample of the 1987-88 NFCS contains data on about 8,500 individuals (with three days of intake data), only a fraction of whom are school children. Second, the SNDA database will contain a wealth of information about the school and the school food service that is not available from the 1987-88 NFCS.²⁰

The SNDA database is unlikely to be an important source of data for FNS' *general* research needs. This study focuses on a single subgroup of the population (school-age children) and so cannot be used for any research that requires estimates for the entire low-income population in the United States.

²⁰The School Nutrition Programs Update survey collected information about school districts that included children who responded to the 1987-88 NFCS. The survey collects information on meal prices, but does not contain detailed information on individual schools.

V. FNS' NEEDS FOR SURVEY DATA: OPTIONS ON HOW TO MEET EACH NEED

This chapter describes the options open to FNS to meet its needs for survey data on food consumption. Chapter II described general areas of research and some specific research projects planned by FNS that would have used the 1987-88 NFCS if the response rate had been higher. We organize this chapter around those same general areas of research, and for each research project that falls within those areas, we summarize its data needs and describe the problems of using the 1987-88 NFCS to meet those needs. We also discuss alternative databases for each project in terms of the advantages that make them suitable for, and the disadvantages that prevent them from meeting the data needs of that project. Our recommendations on how FNS should meet its data needs are discussed in the next chapter. Table V.1 summarizes by research project the problems of using the 1987-88 NFCS and lists alternative sources of data.

The research projects around which this chapter is organized are ones in which FNS has expressed at least some interest. Indeed, FNS has already awarded contracts to conduct most of these projects, but that work is on hold pending the decision regarding the appropriateness of using the 1987-88 NFCS. Several of the projects are only at the discussion or design stages at FNS. Omitted from this chapter are projects that independent researchers might wish to conduct on the basis of food consumption data. These projects are outside the contractual scope of work for this study, which was intended to assess FNS' needs for data on food consumption. Such projects might well have data needs that differ from those of the projects on FNS' research agenda.

A. EVALUATING THE IMPACT OF FNS' PROGRAMS ON FOOD CONSUMPTION

This section discusses possible ways in which FNS could meet its need for data to evaluate the impact of USDA food assistance programs on food consumption. Both the data needs and the options on how to meet those needs vary by the program being evaluated. Evaluations of the impact of the FSP on food consumption of the type that FNS has typically sponsored in the past and has

TABLE V.1

**EXAMPLES OF FNS' RESEARCH PROJECTS, PROBLEMS OF USING THE 1987-88 NFCS
FOR THOSE PROJECTS, AND ALTERNATIVE SOURCES OF DATA**

Research Project	Major Problems of Using the 1987-88 NFCS	Alternative Sources of Data¹
A. Evaluating the Impact of FNS' Programs on Food Consumption		
1. Estimating the Impact of the FSP on Household Food Use	<ul style="list-style-type: none"> • Estimates may be biased 	<ul style="list-style-type: none"> • CES • Low-income sample of the 1977-78 NFCS • 1979-80 SFC-LI • Food Stamp Cashout Demonstration Evaluations • New household food use survey of low-income households
2. Estimating the Impact of WIC on the Diets of Women, Infants, and Children	<ul style="list-style-type: none"> • Sample is too small to make separate estimates of the impact of WIC on infants, pregnant women, breastfeeding women, and postpartum women • Estimates may be biased 	<ul style="list-style-type: none"> • Combined basic samples of 1989-91 CSFII • NHANES III
3. Evaluating the Effectiveness of Nutrition Education Programs	<ul style="list-style-type: none"> • Contains no information on knowledge and attitudes about diet • Contains no information on participation in nutrition education programs • Neither a longitudinal database, a series of cross-sectional databases, nor a specialized cross-sectional database 	<ul style="list-style-type: none"> • DHKS linked to CSFII for 1989, 1990, and 1991 • New specialized survey
B. Program Design		
1. Revising the Thrifty Food Plan	<ul style="list-style-type: none"> • Sample may be too small to estimate model • Sample households may not be representative of all low-income households 	<ul style="list-style-type: none"> • Low-income sample of the 1977-78 NFCS • New household food use survey of low-income households; intake data from combined low-income samples of 1989-91 CSFII or NHANES III
2. Estimating Economies-of-Scale Adjustments in FSP Benefits	<ul style="list-style-type: none"> • Sample households may not be representative of all low-income households 	<ul style="list-style-type: none"> • New household food use survey of low-income households
3. Changing (1) the FSP Benefit Issuance System, (2) Criteria for Authorizing Retail Stores, and (3) Coupon Denominations	<ul style="list-style-type: none"> • Does not contain detailed information on food purchases 	<ul style="list-style-type: none"> • Retail store checkout scanner data

TABLE V.1 (continued)

Research Project	Major Problems of Using the 1987-88 NFCS	Alternative Sources of Data ¹
C. Modeling Program Participation		
1. Modeling Participation in the NSLP and SBP	<ul style="list-style-type: none"> • Sample children may not be representative of all eligible school children • Sample may be too small to estimate models 	<ul style="list-style-type: none"> • SIPP • SNDA Study
D. Descriptive Studies of Food Consumption		
1. Describing the Food Choices of Low-Income Households	<ul style="list-style-type: none"> • Sample households may not be representative of all low-income U.S. households 	<ul style="list-style-type: none"> • Food Stamp Cashout Demonstration Evaluations • New household food use survey of low-income households
2. Describing the Food Intake of Low-Income Children	<ul style="list-style-type: none"> • Sample children may not be representative of all low-income children • Sample is too small for detailed subgroup analyses 	<ul style="list-style-type: none"> • SNDA Study • NHANES III • Combined basic samples of 1989-91 CSFII
E. Methodological Studies		
1. Estimating the Distribution of Usual Dietary Intake with Only One Day of Dietary Intake	<ul style="list-style-type: none"> • Sample is too small to estimate adjustment factors for population subgroups • Adjustment factors may be biased 	<ul style="list-style-type: none"> • Combined basic or low-income samples of 1989-91 CSFII (for estimation of NRC adjustment factors)
2. Developing Methodologies for Analyzing Self-Selection in Models of Multiple Program Participation	<ul style="list-style-type: none"> • Sample may be too small to estimate models 	<ul style="list-style-type: none"> • CSFII • SIPP • NHANES • Food Stamp Cashout Evaluations • 1977-78 NFCS • 1979-80 SFC-LI
3. Validating Data from the Food Stamp Cashout Demonstration Evaluations	<ul style="list-style-type: none"> • Sample households may not be representative of all food stamp households 	<ul style="list-style-type: none"> • CSFII • SIPP • PSID • New household food use survey of low-income households

¹This column provides suggested alternatives to the 1987-88 NFCS. The databases listed in this column are not necessarily recommended to FNS for these projects. We discuss our recommendations in Chapter VI.

several current contracts to conduct require a database of households with income no greater than 130 percent of poverty and data on household food use.¹ While the low-income sample of the 1987-88 NFCS meets both of these requirements, estimates made from this sample may suffer from nonresponse bias. Unfortunately, no other current database contains household food use data on a nationally representative sample of low-income households. Databases created as a result of the Food Stamp Cashout Evaluations contain household food use data, but the samples in these evaluations are not representative of all low-income households. Hence, a research project designed to evaluate the impact of the FSP on household food use would require a new database.

Evaluations of the WIC and school nutrition programs require individual food intake data and a sample of persons from households with income no greater than 185 percent of poverty. As designed, the basic sample of the 1987-88 NFCS could be used for this analysis. However, as the sample is smaller than planned, precise estimates of the impact of these programs on certain subgroups of interest to FNS cannot be made. In addition, estimates may suffer from nonresponse bias. However, unlike the FSP, the impact of WIC and the school nutrition programs could be evaluated using a number of existing surveys, including the 1989-91 CSFII and NHANES III. The School Nutrition Dietary Assessment (SNDA) Study is an additional source of data for analyses of the effects of school nutrition programs on dietary intake.

In the rest of this section, we discuss the FSP, WIC, and nutrition education programs in terms of the data needed to evaluate their impact on food consumption and possible sources for that data.

¹Most of FNS' past or current evaluations of the effects of the FSP on food consumption have entailed the analysis of data on household food use. Examples include the Puerto Rico Nutrition Evaluation (Beebout et al., 1985), the current evaluations of the Food Stamp Cashout Demonstrations in three sites, an econometric analysis of the effects of the FSP on nutrient availability (Devaney, Haines, and Moffitt, 1989), and an econometric analysis of the effects of food stamps on the money value of food used by U.S. households (Devaney and Fraker, 1989). We are aware of only one FNS-sponsored evaluation of the FSP that was based on individual food intake data, the evaluation of the SSI/Elderly Cashout Demonstration (Blanchard et al., 1982).

1. Estimating the Impact of the FSP on Household Food Use

A database used to estimate the impact of the FSP on household food use should have the following five characteristics:

1. It must contain a sample of households with gross income not exceeding 130 percent of poverty.
2. It must contain household food use data.
3. It must contain information on FSP participation, including the value of food stamp benefits.
4. It must contain information on household characteristics, including income; location; and the sex, age, and race of the survey respondent.
5. It must contain information on the proportion of meals eaten by members of the household away from home (to calculate the number of equivalent nutrition units in each household).

a. The 1987-88 NFCS

The 1987-88 NFCS has the five preceding characteristics. Although smaller than planned, the low-income sample of the 1987-88 NFCS contains over 2,500 households with monthly income not exceeding 130 percent of poverty. This is probably large enough to estimate the impact of the FSP on household food use. It is also sufficient to detect, using a one-tailed test, a difference between FSP participants and nonparticipants in the value of purchased food used at home between 5 and 7.5 percent (depending on the value of the mean and standard deviation) and a difference in food energy as a percentage of the RDA available from food used at home of 5 percent (refer back to Table II.2). However, it should be kept in mind that if the weighting procedures led to a serious loss of precision, we may not be able to detect differences this small.

The major problem with using the 1987-88 NFCS to estimate the impact of the FSP on household food use is the potential for nonresponse bias. Estimates made from the 1987-88 NFCS could be severely biased if respondents differed from nonrespondents in unmeasured ways that are correlated with how households use their food stamps. An example better explains the problem.

Suppose that households who care about nutrition are more likely to respond to the 1987-88 NFCS and are also more likely to use additional food stamps to purchase foods rich in vitamin C. If this were the case, estimates from the 1987-88 NFCS would overestimate the impact of the FSP on the availability of vitamin C. The bias arises because the survey respondents are not representative of all FSP households but, in our example, are more likely than nonrespondents to use their food stamps to purchase foods rich in vitamin C.

b. Alternative Sources of Data

We can identify four sources of existing data that could theoretically be used to estimate the impact of the FSP on household food use: (1) the CES, (2) the low-income sample of the 1977-78 NFCS, (3) the 1979-80 SFC-LI, and (4) the Food Stamp Cashout Evaluation databases.

Scarce and Jensen (1979) used the expenditure data from the 1972-73 CES to examine the impact of FSP participation on household nutrient availability. Because of the poor quality of the quantity data in the CES, Scarce and Jensen used the food expenditure data, which were converted to information on nutrient availability in two steps. First, the expenditure data were converted to quantity data by dividing the amount spent on each food by a monthly regional price index for that food. Second, these quantity data were converted to data on nutrient availability using USDA conversion factors.

There are three serious problems with using the CES expenditure data in this way. The first, and most serious, problem is that some food items purchased by households need to be excluded from the analysis. This is because some food categories are so broad and heterogeneous (for example, "refrigerated bakery products") that there is no corresponding regional price index that can be used to convert expenditures to quantities. To the extent that consumption of these excluded food items may differ between FSP participants and nonparticipants, these exclusions may bias the estimates. The second problem is that estimates of nutrient availability will be inaccurate because the specific foods purchased within each food category, and hence nutrient availability, will vary by household.

The third problem is that the analysis assumes that all households in the same region face similar prices. In reality, prices differ depending on location within a region, the types of stores in different neighborhoods, and the household's access to stores. It has been suggested that FSP households face higher than average prices for food. If this is the case, the assumption of uniform prices within a region would result in overestimates of food use by FSP households.

Better alternatives than the CES are the 1977-78 NFCS and the 1979-80 SFC-LI. Both are large enough for this analysis. Both the low-income sample of the 1977-78 NFCS and the 1979-80 SFC-LI contain a sufficiently large sample of low-income households to detect 5 percent differences between food used at home and the availability of energy for FSP households and non-FSP households using a one-tailed test (refer back to Table II.2). Both databases also have all five characteristics required to estimate the impact of FSP participation on household food use. The rate of response to both these surveys was fairly high, suggesting that nonresponse bias is not a serious problem. However, the major disadvantage of using these databases is that they are now 12 to 15 years old. Since the late 1970s, there have been changes in the FSP, in income and prices, new foods have been introduced, and there may have been changes in households' preferences. Thus, the information provided by these databases could be misleading, especially if used as a basis for making policy recommendations.

The fourth alternative source is the data collected in 1990 as part of the Food Stamp Cashout Evaluations that took place in Alabama, San Diego County, and Washington State. These data are recent and contain information on household food use and characteristics as well as FSP participation. Each database contains between 1,000 and 2,300 FSP-participating households. However, about half of these households received food stamp benefits in the form of checks and should be excluded from an evaluation of the FSP as it is currently operated. The database from each demonstration could be merged to create a database containing about 2,300 households that received food stamp *coupons*.

There are two major disadvantages of using the Cashout databases. First, the surveys were conducted in only three states. Hence, to the extent that behavior of households vary across the country, care should be exercised in generalizing the results from these states to the rest of the United States. Second, the surveys did not collect data on FSP-eligible households that did not participate in the FSP. Hence, these data can not be used to examine differences between the food use of FSP households and that of non-FSP households. They can be used, however, to estimate the impact of differences in the size of the FSP benefit on household food use (the "marginal propensity to consume food" out of food stamp benefits).

Apart from the NFCS, the SFC-LI, and the Food Stamp Cashout Evaluation databases, no other available database contains information on household food use. Thus, the only other option open to FNS would be to conduct a new survey to collect household food use data from low-income households. Such a survey would need to collect household food use data on a nationally representative sample of about 3,000 low-income households if it is to be used to detect differences of about 5 percent in food use by FSP and non-FSP households. It could be similar to the 1979-80 SFC-LI except that it would not need to collect any individual intake data. The advantage of a database generated by a new survey would be that it could be specifically designed to meet all five requirements needed to estimate the impact of the FSP on household food use. Its disadvantages would be that the survey would be costly to field, and the data would not be available to analysts for several years. We discuss in more detail in Chapter VI the advantages and disadvantages of conducting a new survey.

2. Estimating the Impact of WIC on the Diets of Women, Infants, and Children

A database used to estimate the impact of WIC on the diets of women, infants, and children should have the following four characteristics:

1. It should contain a sample of WIC-eligible individuals who reside in households with income not exceeding 185 percent of poverty.

2. It should contain data on food and nutrient intakes by individuals. More than one day of intake data for some individuals in the sample is required to estimate the distribution of usual nutrient intake.
3. It should identify which persons are WIC participants and whether individuals satisfy the WIC income and categorical eligibility criteria (that is, they are children younger than 5 or women who are pregnant, breastfeeding, or postpartum and are in households with income less than 185 percent of poverty).²
4. It should include such household characteristics as income; location; and sex, age, and race of the household head.

a. The 1987-88 NFCS

The 1987-88 NFCS has the four preceding characteristics. However, the high rate of nonresponse to the 1987-88 NFCS creates two serious problems in using the database to estimate the impact of WIC on food and nutrient intake. First, as both the basic and low-income samples of the survey are smaller than planned, the database may not be large enough to make precise estimates of the impact of WIC. Using a one-tailed test to detect a 7.5 percent difference in the intake of food energy between women who participate in WIC and those who do not requires a sample of over 500 WIC-eligible women (refer back to Table II.3).³ A sample of over 300 WIC-eligible children is required to detect a 7.5 percent difference in the intake of food energy between WIC children and non-WIC children (refer back to Table II.3). While there are a sufficient number of WIC-eligible children in the basic sample of the 1987-88 NFCS to detect a 7.5 percent difference in food energy intake, there are only 62 women in the basic sample who are in households with income not exceeding 185 percent of poverty and who are either pregnant or breastfeeding. With this number of WIC-eligible women, the analyst would not be able to detect any difference in energy intake that is smaller than about 20 percent.

²Individuals also have to show they are nutritionally at risk to be eligible for the WIC program.

³For expositional convenience, we use the term "WIC eligible" to refer to individuals who meet the WIC income and categorical eligibility criteria but may or may not satisfy the nutritional risk criteria.

While the number of WIC-eligible women and children is probably larger in the low-income sample of the NFCS than in the basic sample, the low-income sample does not contain any households with income over 130 percent of poverty.⁴ This is a problem because in many states the income-eligibility threshold for WIC is higher than 130 percent, and in some, it is as high as 185 percent of poverty. Therefore, estimates of the impact of WIC made using the low-income sample of the NFCS should not be generalized to the whole WIC-eligible population.

One solution to the problem of finding a database to estimate the impact of WIC would be to merge the low-income and basic samples of the 1987-88 NFCS. The merged sample could be used to estimate the impact of WIC on food intake; it may also contain enough observations to estimate the impact of WIC on women. However, it is doubtful whether it would yield enough observations to estimate separately the impact of WIC on pregnant, breastfeeding, and postpartum women. As discussed in Chapter III, merging the low-income and basic samples of the 1987-88 NFCS would require substantial input from high-level technical staff, and although it would address the problem of small sample size, it would not solve the problem of nonresponse bias.

Nonresponse bias would occur if WIC-eligible individuals who responded to the survey are not representative of all WIC-eligible individuals, and if the decision to respond to the survey is correlated with the impact of the WIC program on participants' diets. For example, suppose that WIC has the greatest impact on the diet of pregnant women who are already receiving prenatal care, and that women receiving prenatal care were more likely to respond to the NFCS. In this example, the estimates of the impact of WIC on food and nutrient intake would overestimate the true impact of the WIC program.

⁴As data on individuals in the low-income sample of the 1987-88 NFCS have not yet been released, we do not know the number of WIC-eligible women and children in the low-income sample.

b. Alternative Sources of Data

Two databases other than the 1987-88 NFCS could be used to estimate the impact of WIC on diet: (1) the CSFII, and (2) NHANES III. The CSFII includes all the information needed to estimate the impact of WIC on diet. The major advantage of using the CSFII is that it collects more than one day of food intake data from most individuals in the sample. The major disadvantage is that its samples are relatively small. The basic sample of the 1989 CSFII contains at least one day of food intake data on 3,502 individuals. If the proportion of WIC-eligible individuals in the basic sample of the 1989 CSFII is the same as the proportion in the basic sample of the 1987-88 NFCS, the 1989 CSFII will contain only about 20 WIC-eligible women and 140 WIC-eligible children. Like the 1987-88 NFCS, the low-income sample of the 1989 CSFII may contain more WIC-eligible women and children, but it does not contain any households with income exceeding 130 percent of poverty. Hence, because it does not include WIC-eligible persons in households with income between 130 and 185 percent of poverty, estimates using this sample should not be generalized to the entire WIC-eligible population.

Merging the basic samples of the 1989, 1990, and 1991 CSFII would be one way to take advantage of the database. This is a relatively simple procedure that would increase the sample size by a factor of three (if the response rates to the 1990 and 1991 CSFII are similar to the response rates of the 1989 CSFII). The resulting sample of WIC-eligible children would be large enough to detect a 7.5 percent difference between the food energy intake of WIC-children and that of non-WIC children. However, the sample of women would still be too small to estimate the impact of WIC on women. More precise estimates of the impact of WIC could be obtained by merging the basic and low-income samples of each CSFII in addition to merging across all three years. However, as we discussed in Chapter IV, merging two different samples would probably yield a combined sample large enough to estimate the impacts of WIC on eligible children and women, but not large enough to estimate the separate impacts of WIC on pregnant, breastfeeding, and postpartum women.

NHANES III could be used to estimate the impact of WIC since it contains information on WIC participation and enough information to determine whether each person is income and categorically-eligible for WIC. This database is extremely large: it is projected to contain information on about 35,000 persons. Even though NHANES does not oversample low-income persons, there would be enough WIC-eligible persons in the sample to yield precise estimates of the impact of WIC on women and children, and possibly on pregnant, breastfeeding, and postpartum women. Another advantage of NHANES is that data from the first three years of the survey will be available in 1992. A disadvantage is that it collects only one day of food intake data. Hence, it would not be possible to make unbiased estimates of the proportion of the WIC-eligible individuals who are nutritionally at risk from estimates of the distribution of usual intake. It would, however, be possible to obtain unbiased estimates of the mean usual intake for various subgroups targeted by WIC.

3. Evaluating the Effectiveness of Nutrition Education Programs

A database used to evaluate the effectiveness of a nutrition education program should contain information on (1) participation in (or exposure to) the nutrition education program, (2) a measure of food consumption, and (3) people's knowledge and attitudes about diet and nutrition. In order to compare the diet and diet-related knowledge and attitudes of people who participate in the program with that of people who do not participate in the program, the database should contain sufficient numbers of both program participants and nonparticipants. If the program is introduced nationwide, thus potentially affecting the entire population, the database should be longitudinal or contain a series of independent cross-sections so that diet and diet-related knowledge and attitudes before the introduction of the program can be compared with diet and diet-related knowledge and attitudes after the introduction of the program.

a. The 1987-88 NFCS

The 1987-88 NFCS is not suitable for evaluating the impact of nutrition education programs. Although it contains a measure of food consumption, it provides information neither about participation in nutrition education programs nor about diet-related knowledge and attitudes.⁵ Moreover, it could not be used to evaluate a small-scale nutrition education program because it would not contain a sufficient number of program participants. And because it is not a longitudinal database, it is not appropriate for evaluating the impact of a nationwide nutrition education program.

b. Alternative Sources of Data

The DHKS--a follow-up survey to the CSFII--is the only large-scale survey of people's knowledge and attitudes about diet and nutrition. Linked to the CSFII, it also provides data on individual food intake. As a series of similarly formatted cross-sectional databases, the 1989, 1990, and 1991 DHKS could be used to evaluate a nationwide nutrition education program if one had been introduced or expanded during that period.

Despite these advantages, the DHKS is not suitable for evaluating smaller-scale nutrition education programs. At present, it does not collect any information about participation in nutrition education programs. Even if it was revised to do so, it is unlikely that the DHKS would contain enough participants to evaluate the effectiveness of a program, unless the program affected a large proportion of the population.

No current database is suitable for evaluating small-scale nutrition programs. For this purpose, a special survey would be necessary. It should contain questions about participation in nutrition education programs in addition to the questions about diet, and about diet-related knowledge and attitudes that are contained in the CSFII and DHKS. It should be administered to enough program participants and program-eligible nonparticipants that live in the vicinity of the program site to enable

⁵The 1987-88 NFCS does include a question on sources of nutrition information (for example, relatives or friends, food packages or labels, and so on). This is not sufficient information on which to base a study of the effectiveness of nutrition education programs.

the analyst to detect differences between participants and nonparticipants in terms of diet-related knowledge, attitudes, and behavior.

B. PROGRAM DESIGN

In this section, we discuss possible ways for FNS to meet its data needs for evaluating and potentially changing the design of the FSP. We consider the same three changes discussed in Chapter II: (1) revising the TFP, (2) revising the economies-of-scale adjustments in the benefit schedule, and (3) other changes in the design of the FSP, such as changing the method by which benefits are issued, changing the retail store authorization policy, and changing the denominations of food stamp coupons.

1. Revising the Thrifty Food Plan

A database used to revise the quantities of food that make up the TFP should meet the following three requirements:

1. It should contain an up-to-date, large sample of nationally representative low-income individuals.
2. It should contain information on the types and quantities of food used by individuals *or* on individual food intake, and the types and quantities of food used by households.
3. It should contain information on the prices of food purchased for use at home.

a. The 1987-88 NFCS

Although the low-income sample of the 1987-88 NFCS partially meets these requirements, there are two major problems in using the 1987-88 NFCS to revise the TFP quantities. First, a revision of the TFP requires enough observations to estimate the intake of food from each of the 31 TFP food groups by individuals in each of the 11 TFP age-sex categories, and the sample may not be large enough to obtain precise estimates of the quantities of food used in each of these subgroups. The low-income sample of the 1977-78 NFCS, used in the last revision of the TFP, contained about 4,600

households. The low-income sample of the 1987-88 NFCS contains only about 2,500 households, just over half the number of households in the sample previously used to revise the TFP. Whether the low-income sample of the 1987-88 NFCS is sufficiently large requires further investigation.

A second major problem of using the 1987-88 NFCS is that the high rate of nonresponse casts doubt on whether the sample of low-income households is truly representative of all low-income households. As the TFP is used to determine the food stamp benefit schedule, with direct implications for the well-being of FSP households and program costs, it is extremely important that it is based on household food use data that are representative of the food used by all low-income households.

b. Alternative Sources of Data

Other than the 1987-88 NFCS, no current database has all three characteristics required to revise the quantities in the TFP. The low-income sample of the 1977-78 NFCS does have all the necessary characteristics, but it is now over a decade old. One option would be for HNIS to revise the TFP using the low-income sample of the 1977-78 NFCS, along with the most recent RDAs and dietary guidelines and the most recent factors from the National Nutrient Data Bank to convert quantities of food into their component nutrients. Thus revised, the TFP would improve upon the current TFP in that it would reflect changes in the RDAs and dietary guidelines, changes in the nutrient content of foods, and changes in the factors that are used to convert food into nutrients. However, such a revision would not account for possible changes in relative food prices or in the food choices of low-income individuals that may have occurred since the fielding of the 1977-78 NFCS.

A complete revision of the quantities that make up the TFP would require the fielding of a new survey that would collect data on household food use from a nationally representative sample of low-income households. A sample of about 4,500 to 5,000 households would be needed to obtain estimates that are about as precise as those made in the 1983 revision of the TFP.

Ideally, this survey should also contain individual food intake data, but this would add to the burden on respondents and the cost of the survey. However, if individual food intake data are available from an alternative source, this would significantly reduce the respondent burden and cost of the survey. Individual food intake data can be used to estimate the proportion of the intake of each type of food that can be accounted for by individuals in each of the TFP age-sex categories. These proportions can then be used to convert data on household food use to data on individual food use. As these proportions are not household specific, they can be estimated using one database and applied to household food use data from another. Either the 1989-91 CSFII (merged across years) or NHANES could be used to estimate these proportions. This is a logical extension of the methodology that was applied to the 1977-78 NFCS to obtain the current TFP.

2. Estimating Economies-of-Scale Adjustments in FSP Benefits

Estimating economies-of-scale adjustments to be used in determining food stamp benefit allotments for households of different sizes requires a database with the following three characteristics:

1. It should contain a nationally representative sample of low-income households.
2. It should contain data on the money value and quantity of food used by the household.
3. It should contain data on household composition, household income, the value of any food stamps received by the household, and the number of meals eaten away from home by household members. If the price-variation approach to modeling economies of scale is used (see Chapter II), it should also contain data on prices.

a. The 1987-88 NFCS

As designed, the low-income sample of the 1987-88 NFCS would have been particularly well-suited for estimating economies of scale. It was to contain about 3,600 low-income households, and it contains information on both the money value and quantity of food used by the household. However, given its high rate of nonresponse, the 1987-88 NFCS may not be suitable for estimating

economies of scale for at least two reasons. First, the households that responded to the 1987-88 NFCS may not be representative of all low-income households. It is particularly worrisome that two-adult households were found to be overrepresented in the sample (LSRO, 1991), which suggests that household size may have affected the decision to respond to the survey. Second, and perhaps less important, the actual size of the low-income sample may be too small to make precise estimates of the impact of household size on per person food expenditures.

b. Alternative Sources of Data

The only databases other than the NFCS that contain household food use data are the Food Stamp Cashout Evaluation databases. While these databases contain information on the money value and quantity of food used at home, and information on food stamp benefits and the composition of the household, estimates of economies of scale developed from these databases should not be used as a basis for altering the current economies-of-scale adjustments used in the FSP for two reasons. First, households in these databases are not nationally representative. Second, all the households in the databases participate in the FSP and so are not representative of all low-income households. Therefore, the only other option would be to conduct a new survey that would collect data on the money value and quantity of food used by the household from a nationally representative sample of low-income households. A sample of about 4,000 households was sufficient to estimate the current economies-of-scale adjustments and would probably be sufficient to revise these estimates.

3. Changing the FSP Benefit Issuance System, Criteria for Authorizing Retail Stores, and Coupon Denominations

Changing the design of the FSP may involve changes in how benefits are issued, how retail stores are authorized to accept coupons, and in coupon denominations. The development of such design changes requires information about the food purchasing behavior of food stamp recipients and the following other kinds of information:

- Quantity and type of food purchased
- Form in which the benefit is issued (coupon, check, or EBT)
- Form in which the payment for food is made (coupon, cash, or EBT)
- Time during the month the food purchases are made
- Type of store at which the purchase is made
- Size of each transaction
- Value of coupons used and amount of change provided

a. The 1987-88 NFCS

The NFCS does not contain enough information about the food purchasing behavior of food stamp recipients to be used to design FSP changes of these types.

b. Alternative Sources of Data

No existing database contains sufficient information about food stamp recipients' food purchasing behavior to address all possible program design changes. However, the data collected by retail food stores via their scanner systems has the potential to provide this information. Scanner systems collect data on the UPC code (which identifies a product and its size) and the price of each purchased item that has a UPC code. As explained in Chapter IV, variable-weight foods, which do not currently have UPC codes, account for about 40 percent of all retail food sales. Scanner systems also record the amount of each transaction paid for by food stamps and the amount paid for by cash.

It is not the current practice of retail food stores to maintain data on individual transactions; rather, they typically maintain sales data aggregated across customers and over time by type of product. Recently, some stores have begun to maintain separate data files on purchases made by individual "preferred customers." However, even these data are aggregated over time by type of product, and hence, information on individual transactions is lost. These files also include basic income and demographic information provided by the customer. The absence of information on

individual transactions in either of these types of data files limits their usefulness for evaluating the current design of the FSP. However, the rapid pace of change in the storage and use of data on retail food sales suggests that this may be a valuable source of data for FNS within a few years.

Information on the characteristics of households that make the food purchases together with data on the stores at which food is purchased and the types, quantities, and prices of foods purchased is commercially available (for example, from A.C. Nielsen or Information Resources, Inc.). However, no information is collected on whether the household participates in the FSP or whether the household uses food stamps to purchase specific food items. It may be possible for FNS to negotiate the collection of these types of data.

The linkage of EBT systems and scanner systems in retail food stores in EBT demonstration sites could provide a wealth of data on the food purchases of food stamp households. However, such links do not currently exist, or in those demonstration sites where they do exist, considerable additional systems development would be required before data useful to FNS could be extracted. Because EBT eliminates the need for food stamp coupons, linked EBT and scanner systems have no potential for providing data on the use of coupons.

C. MODELING PROGRAM PARTICIPATION

In this section, we discuss FNS' needs for data to model participation in USDA food assistance programs and the options to meet those needs. As in the other areas of research, the types of data needed to model participation in food assistance programs vary by program. However, in most instances, a database must meet three requirements if it is to be used to model program participation:

1. It should be based on a large, nationally representative sample containing households (or individuals) that are eligible for the program.
2. It should contain sufficiently detailed information about the characteristics of the households (or individuals) to approximate their eligibility for the program.
3. It should contain information on whether the household (or individual) participates in the program of interest.

As it was designed, the 1987-88 NFCS meets these three requirements for modeling participation in the FSP. However, even as designed, the sample sizes of the 1987-88 NFCS may not be large enough for some modeling of participation in the WIC and school nutrition programs. This is because both of these programs serve relatively small subgroups of the population. In addition, because the income cutoff for the *low-income sample* of the NFCS is 130 percent of poverty, and the income cutoff for the WIC and school nutrition programs is 185 percent of poverty, the low-income sample alone cannot be used to model participation by all eligibles in these programs. In contrast, the *basic sample* of the 1987-88 NFCS does contain households with incomes between 130 and 185 percent of poverty. However, because low-income households are not overrepresented in the basic sample, there may be too few of them in this sample to model participation in the WIC and school nutrition programs.

This problem of sample size is exacerbated by the high rate of nonresponse in the 1987-88 NFCS. The high rate of nonresponse also resulted in a sample that may not be representative of all U.S. households, a problem that is especially serious if the decision to respond to the NFCS is correlated with the decision to participate in the program.

Given the problems with the 1987-88 NFCS, it may be preferable to use other existing nationally representative databases to model program participation. SIPP contains detailed program participation and household characteristic data on over 20,000 households. Alternatively, NHANES contains some information on participation in the FSP, WIC, and the school nutrition programs for about 35,000 individuals. More detailed program data on participation in school nutrition programs can be obtained from the database for the School Nutrition Dietary Assessment Study.

Most of the existing literature on the modeling of participation in USDA food assistance programs has focused on the FSP.⁶ This has also been the focus of most FNS-sponsored research on program participation (Czajka, 1981; Fraker and Moffitt, 1988; Martini, 1992). Nevertheless, we

⁶Allin and Beebout (1989) review the literature on the modeling of participation in the FSP.

have selected a study of participation in the SBP and NSLP to illustrate FNS' needs for data to model program participation. We have done so because FNS had firm plans to update its existing models of participation in these programs and, to facilitate that work, had undertaken a supplemental survey of the prices charged for USDA breakfasts and lunches by the school districts in which children in the 1987-88 NFCS database reside.

1. Modeling Participation in the NSLP and SBP

Modeling participating in the NSLP and SBP requires a nationally representative sample of school-age children from households whose income does not exceed 185 percent of poverty. Ideally, the database should contain information on:

- Whether the child attends a school that participates in the NSLP and SBP
- How frequently a child eats school lunches or breakfasts
- The full and reduced price of the meals
- The meal-price status of the child (that is, whether the child can receive free or reduced-price meals)

a. The 1987-88 NFCS

The 1987-88 NFCS asked four questions about participation in both the school lunch and school breakfast programs:

1. Does the child attend a school that serves a school lunch or breakfast?
2. Approximately how many times a week does the child usually eat a school lunch or breakfast?
3. If a child eats school lunches or breakfasts, does the child get them free?
4. If a child eats school lunches or breakfasts, how much does the child pay for them?

The answers to these questions along with other variables in the NFCS provide nearly all of the information needed to model participation in the school nutrition programs. The first two questions

on the NFCS about school nutrition programs provide information on whether the child attends a school that participates in the school nutrition programs and the frequency with which the child eats school meals. The meal-price status of most children can be inferred from the third and fourth questions about school meals on the NFCS. For those children who eat school meals, the NFCS tells us whether they receive free meals and the price they pay if the meal is not free. If we know that a child receives a meal, the price the child pays for the meal can be used to infer whether it is a reduced-price meal. However, if a child does *not* eat school meals, we cannot infer from the information provided by the NFCS the price that would have been paid if a child had eaten the meals.

While the NFCS does not give us the full price of a meal for all children, the School Nutrition Programs Update database does contain information on school meals, including price, from all school districts attended by children who responded (or whose parents responded) to the 1987-88 NFCS. Because this database has been linked to the NFCS, it is not difficult to use it to obtain the full and reduced prices of the meals available to the children in the sample.

The 1987-88 NFCS, together with the School Nutrition Programs Update database, provides data on most of the variables required to model participation in the school nutrition programs. However, the high rate of nonresponse to the 1987-88 NFCS creates two problems that make it unsuitable for modeling participation in the NSLP and SBP. First, the households in the basic sample of the 1987-88 NFCS may not be representative of all U.S. households. Thus, it may be dangerous to use models estimated on the basis of the 1987-88 NFCS to simulate the impact of program changes on program cost and the number of people who receive benefits. Second, the basic sample of the 1987-88 NFCS may not be large enough to provide precise estimates of the determinants of participation. The basic sample of the 1987-88 NFCS contains at least one day of food intake data for only about 840 children who both attend a school that provides lunches and who reside in a household that is income eligible for the NSLP, and only about 340 children who both attend a school that provides breakfasts and

who reside in a household which is income eligible for the SBP (refer back to Table III.3). Once these children are grouped by their meal-price status, the subgroups could become very small. The low-income sample of the NFCS cannot be used for modeling participation in the school nutrition programs because it does not contain households that are income-eligible for reduced-price meals but not for free meals, that is, households with income greater than 130 and less than 185 percent of poverty.

b. Alternative Sources of Data

One alternative to using the 1987-88 NFCS to model participation in the school nutrition programs is SIPP. SIPP contains a large sample of households and information on participation in the NSLP and SBP, including (1) the number of children in the household who receive free or reduced-price school lunches or breakfasts and (2) the child's meal-price status. Unfortunately, SIPP does not contain any information on how frequently the child purchases school lunches or breakfasts, the price of the meal, or whether the child attends a school that serves breakfasts or lunches. This would severely limit the use of SIPP data as a means of modeling participation in the NSLP or SBP. In addition, SIPP would provide no option for extending a study of participation in school nutrition programs to include an assessment of the dietary impacts of participation.

A better alternative would be the data collected as part of the SNDA Study because it will include data on about 4,000 nationally representative school children. These children attend schools that both participate and do not participate in USDA school nutrition programs. The database will include information on (1) which children eat USDA meals at school, (2) the full and reduced prices of the school meals, and (3) the meal-price status of each child. In addition, it will contain a wealth of information about the household and the school. However, it does not include the frequency with which children eat school meals. Hence, while this database will support the modeling of the child's decision to eat a school meal on any given day, it will not support the modeling of the child's decision

to eat a school meal on any day during a given period.⁷ However, it will be possible to model the school's decision to provide school lunches and breakfasts and the parent's decision to apply for free or reduced-price meals. And as the SNDA database will include dietary intake data, it will offer the option of linking an analysis of participation with one or the dietary impacts of participation.

D. DESCRIPTIVE STUDIES OF FOOD CONSUMPTION

This section discusses the data needs of two descriptive studies of food consumption by groups of households and individuals of interest to FNS. The first is a descriptive study of the food choices of low-income households. The second is a descriptive study of the food intake of low-income children. The data needs and the options for meeting them are unique to each study. FNS has already awarded contracts to conduct these studies. At the time those contracts were awarded, FNS had intended that both studies would be conducted on the basis of data from the 1987-88 NFCS.

1. Describing the Food Choices of Low-Income Households

Describing the food choices of low-income households requires a sufficiently large sample of low-income households, and information on their food use, nutrient availability, and participation in food assistance programs.

a. The 1987-88 NFCS

If the response rate had been higher, the low-income sample of the 1987-88 NFCS would have met the data needs of this project. For this type of study, the problem of nonresponse bias is much greater than the problem of sample size. Although the sample size is smaller than planned, it is not much smaller than the sample used successfully in a previous study of food choices of low-income households using the 1979-80 SFC-LI (Kisker and Devaney, 1988). Hence, for this application the small sample size arising from nonresponse is not a serious issue.

⁷For example, the SNDA database could not support the modeling of a child's decision to any USDA school lunches during a week.

However, the potential for nonresponse bias is a serious problem for this type of descriptive study. One of the main purposes of this study is to generalize its results to the U.S. low-income population. If respondents to the 1987-88 NFCS differ from nonrespondents in ways that are correlated with their food choices, it would be incorrect to generalize from this sample to all low-income households. It would be similarly incorrect to attribute differences between the food choices of households in the 1987-88 NFCS and households in earlier surveys to real changes in the household's food choices over time. Any apparent change over time might result solely from differences in the characteristics of households who chose to respond to the surveys.

b. Alternative Sources of Data

Apart from the NFCS, the only other current databases that contain data on household food use are the databases developed from the Food Stamp Cashout Evaluations. If these three databases were merged, the resulting database would be sufficiently large to describe the food choices of low-income households. However, because these databases are not nationally representative, the results of a study based on the Cashout databases could not be generalized to all U.S. low-income households.

The only other option open to FNS would be to use data from a new survey of low-income households. Such a survey would need to collect data on household food use, participation in food assistance programs, and some demographic and socioeconomic characteristics of the households. A sample of about 3,000 households, about the same size as the 1979-80 SFC-LI, would be sufficient.

2. Describing the Food Intake of Low-Income Children

A descriptive study of the food intake of low-income children requires a database containing a sample of low-income children and information on their nutrient intake, participation in USDA food assistance programs, and demographic and socioeconomic characteristics of the household. If the

food intake of low-income children is to be compared with that of higher-income children, the sample should also include higher-income children.

a. The 1987-88 NFCS

It was originally suggested that the 1987-88 NFCS might be a useful source of data on the adequacy of diets of low-income children. Its advantage is that it contains three days of food intake data for at least some of the children in the sample, which can be used to estimate the distribution of usual nutrient intake and the proportion of children that are nutritionally at risk. However, this advantage is offset by two potentially important problems that result from a high rate of nonresponse.

First, the high rate of nonresponse means that children who participated in the 1987-88 NFCS may not be representative of all low-income children. This could give a misleading picture of the food intake of the current dietary status of low-income children. Comparisons between the data for children in this survey with those from earlier surveys may yield especially misleading results if the decision to respond to the survey is correlated with an outcome of interest. Comparisons between different subgroups of children using the 1987-88 NFCS may also yield misleading results if the factors that affect the decision to respond to the survey vary across subgroups. For example, suppose that low-income households are more likely to respond to the survey if they have an adequate diet, but higher-income households respond to the survey at the same rate regardless of the quality of their diet. In this case, the survey might incorrectly lead us to conclude that the quality of children's diets does not improve with income.

Second, because the actual sample sizes of the 1987-88 NFCS are fairly small, the extent to which it can support a descriptive analysis of the characteristics of children in small subgroups of interest to FNS is limited. For example, the 1987-88 NFCS is almost certainly too small to describe the nutrient intake of persons under age 18 who are pregnant or breastfeeding.

b. Alternative Sources of Data

There are three databases that could at least partly substitute for the 1987-88 NFCS: (1) the database collected as part of the SNDA Study, (2) NHANES III, and (3) the 1989-91 CSFII. All three databases collect data on children's food intake. One problem with the SNDA database is that it will not contain data on children of all ages; rather, it will contain data on school-aged children only. Hence, while this database could be used alone to support a study of all low-income children.

NHANES III will be a large database (about 35,000 individuals) that contains information on children of all ages. It also provides data on participation in the FSP, WIC, NSLP, and SBP. Because the database is large, and the survey oversamples children, NHANES may be able to support some fairly detailed subgroup analyses. The major drawback of NHANES and of the SNDA database is that they provide only one day of individual food intake data. Thus, these databases alone cannot be used to estimate the distribution of usual nutrient intake or to estimate the proportion of the population that is nutritionally at risk. However, it would be possible to use these databases in conjunction with data on multiple days of nutrient intake from the 1989-91 CSFII to obtain these estimates. This would entail using a procedure recommended by the National Research Council (1986) to estimate an adjustment factor on the basis of the CSFII data. This adjustment factor could then be applied to the NHANES or SNDA databases to obtain the desired estimates. It would be necessary to estimate a unique adjustment factor for each demographic group of interest. The following section provides a more complete discussion of the adjustment procedure. The only databases other than the 1987-88 NFCS that contain more than one day of individual food intake data are the databases for the 1989-91 CSFII, which contain up to three days of intake data for individuals in the samples. The CSFII databases also contain information on whether the individual participates in USDA food assistance programs, and demographic and socioeconomic characteristics of the household. However, the basic sample of the 1989 CSFII contains fewer than 3,000 individuals

who reported all three days of intake data, and the low-income sample contains fewer than 1,400 such individuals.⁸ These sample sizes are too small to support a reliable subgroup analysis of children.

One option would be to merge the basic samples of the 1989, 1990, and 1991 CSFII. The resultant database would contain about 8,800 individuals and about 2,600 children who reported three days of intake data.⁹ Although this database would probably still not be large enough to support analyses of small subgroups of children, it could be used for some subgroup analysis; and unlike the 1987-88 NFCS, the CSFII does not have the problem of high rates of nonresponse.

E. METHODOLOGICAL STUDIES

In this section, we describe three research projects aimed at improving research techniques and methodologies: (1) estimating the distribution of usual dietary intake with only one day of intake data, (2) developing methodologies for analyzing self-selection in models of multiple program participation, and (3) validating data from the Food Stamp Cashout Evaluations. The extent to which the needs of each of the projects are met by both the 1987-88 NFCS and alternative databases varies by project. For the first and third projects, the possible nonrepresentativeness of the households in the 1987-88 NFCS is a serious problem. For the second project, the relatively small sample size of the 1987-88 NFCS is the more serious drawback.

When estimating the distribution of the usual dietary intake for some populations, the CSFII may be a good substitute for the 1987-88 NFCS. However, there is no good substitute for the 1987-88 NFCS for validating the data from the Cashout Evaluations. On the other hand, several databases could be used to develop econometric and computer models for analyzing self-selection in the presence of multiple program participation.

⁸HNIS provided these preliminary counts of the 1989-91 CSFII sample sizes.

⁹These estimates assume that the sample sizes and the proportion of children in the sample in the 1990 and 1991 CSFIIs are about the same as in the 1989 CSFII.

1. Estimating the Distribution of Usual Dietary Intake with Only One Day of Dietary Intake Data

Because of intraindividual variation in dietary intake, distributions of usual dietary intake developed from databases that contain only one day of dietary intake data from each individual will yield biased estimates of the proportion of the population who are nutritionally at risk (National Research Council, 1986). It has been suggested that databases that contain three days of intake data can be used to adjust the distributions of usual intake estimated from a sample of one-day intake data for the existence of intraindividual variation. This procedure would involve estimating "adjustment factors" from the sample of three-day individual food intake data. To ensure that this procedure is valid, it is important to estimate the adjustment factors from a sample that covers a population similar to that covered by the one-day sample.

a. The 1987-88 NFCS

The 1987-88 NFCS is one of the few databases that collects more than one day of dietary intake from at least some individuals in the sample. Again, the high rate of nonresponse, makes the NFCS unsuitable for estimating adjustment factors for two reasons. First, the small sample size makes it even more difficult to estimate the adjustment factors for a subgroup of the population. For example, it was suggested that the basic sample of the 1987-88 NFCS be used to adjust the distributions of usual dietary intake estimated from the SNDA Study, which will collect only one day of intake data from each child. But the sample of school children in the 1987-88 NFCS basic sample will be too small to develop reliable estimates of the adjustment factors. Second, if the intraindividual variation in the dietary intake of respondents to the survey differs from that of the nonrespondents, the estimates of the adjustment factors from the 1987-88 NFCS will themselves be biased.

b. Alternative Sources of Data

The CSFII is the only other database that contains more than one day of dietary intake data. The advantage of the CSFII over the 1987-88 NFCS is that it may have a higher rate of response (the

response rate for the 1989-91 CSFII is not currently available). However, the CSFII sample size is still a problem. Even if the 1989, 1990, and 1991 basic samples of the CSFII were combined, the resultant database would still contain only about the same number of individuals who provided three days of food intake data as the basic sample of the 1987-88 NFCS. Therefore, it would not be possible to use this database to estimate adjustment factors for certain subgroups of the population.

2. Developing Methodologies for Analyzing Self-Selection in Models of Multiple Program Participation

On at least three occasions, FNS has wished to undertake an analysis of the effect of the FSP on a measure of food consumption or some other outcome and has found that it cannot do so without simultaneously considering the effect of a second closely related program, such as WIC or AFDC (Fraker and Moffitt, 1988; Fraker, Long, and Post, 1990; Long, 1991). On the basis of this experience, FNS determined that it is likely to have future needs to analyze measures of food consumption that are influenced by the decisions of a household or an individual regarding participation in three or more assistance programs, such as the FSP, NSLP, and SBP; or the FSP, AFDC, and WIC. FNS therefore decided to undertake the development of a general model that would permit the estimation of the effects of three or more assistance programs on such an outcome measure, while simultaneously controlling for the observed and unobserved factors that influence the decisions by eligibles to participate in those programs. The development of this model and the related statistical software requires a database that contain measures of food consumption and measures of participation in food assistance programs and other assistance programs. The database must also contain enough observations on low-income households or individuals to permit the reliable estimation of as many as 40 or more parameters in a complicated nonlinear model. It is possible that the sample size requirement may be made even more stringent by the fact that it may be necessary to estimate the model on a narrowly defined subgroup of the population, such as children ages 1 to 5 years, or school-age children.

a. The 1987-88 NFCS

FNS had intended for this model to be developed and estimated on the basis of data from the 1987-88 NFCS. The advantage of the 1987-88 NFCS for this project is that it contains information on participation in the major USDA food assistance programs and provides several alternative measures of food consumption. The disadvantage is its high rate of nonresponse, which could result in biased estimates of the model. Perhaps more importantly, the NFCS sample sizes may be too small to estimate such a complex model on data for a narrowly defined subgroup. However, if it is determined that the samples are large enough, and if the estimates made from the models are cautiously interpreted and limited in their application, the survey could be used to develop this model. The estimates produced by the model should be used to assess the functioning of the software rather than to guide policymakers.

b. Alternative Sources of Data

As the estimates produced by the initial application of this model may be less important than the development of the model and related software, there are a number of databases that could support that developmental research. The CSFII, NHANES III, the 1977-78 NFCS, the 1979-80 SFC-LI, and the Food Stamp Cashout databases all contain data on both food consumption and participation in USDA food assistance programs. The NHANES and the 1977-78 NFCS may be the most appropriate for this research project because they are the largest of the food consumption databases; the basic sample of the 1977-78 NFCS contains nearly 15,000 households, and the first phase of the NHANES III is projected to contain about 35,000 individuals. However, the disadvantage of the 1977-78 NFCS is that any estimates obtained by applying the model to that database would be 14 to 15 years out of date.

3. Validating Data from the Food Stamp Cashout Demonstration Evaluations

Validating data from the Cashout evaluations entails determining the degree to which those data represent the characteristics and food use of food stamp recipients nationwide. Realistically, it may never be possible to generalize the findings of the Cashout demonstrations to the population nationwide with a high degree of confidence. However, the results from two types of comparisons between food stamp *coupon* recipients in the demonstration sites and food stamp recipients nationwide would likely influence informed judgments regarding the generalizability of those findings: (1) a comparison of the household demographic and socioeconomic characteristics of food stamp coupon recipients and (2) a comparison of household food use of food stamp coupon recipients. These comparisons should be made between each of the Cashout databases and a database that contains a nationally representative sample of food stamp households and information on their food use.

a. The 1987-88 NFCS

As designed, the 1987-88 NFCS could have supported both types of comparisons. However, the results of these comparisons will be questionable if they are based on the data actually obtained by that survey. This is because the high rate of nonresponse calls into question the extent to which the FSP households in the 1987-88 NFCS represent FSP households nationwide.

b. Alternative Sources of Data

The CSFII, SIPP, and the PSID are suitable for making the comparison between household demographic and socioeconomic characteristics of food stamp recipients nationwide and food stamp coupon recipients in the Cashout demonstration sites. Of the three databases, SIPP is probably the best choice because it is such a large database, containing information on about 21,000 households, and it also contains a wealth of socioeconomic and demographic data.

No existing database is suitable for making the comparison between the food use of demonstration households and FSP households nationwide. The 1977-78 NFCS and the 1979-80 SFC-LI, the only other databases that contain information on household food use, are too old to allow analyst to distinguish between location-related differences in food use and differences that are purely a result of changing food consumption patterns over time.

One option would be to conduct a new survey that would collect data on food use and socioeconomic and demographic characteristics of FSP households. The sample for this survey should be large enough to allow some subgroup comparisons, such as for urban and rural households.

VI. RECOMMENDATIONS ON WAYS TO MEET FNS' NEEDS FOR SURVEY DATA ON FOOD CONSUMPTION

This chapter presents our recommendations to FNS on how to meet its needs for survey data on food consumption. These recommendations cover the use of the 1987-88 NFCS, the use of alternative databases, the fielding of a new survey of food use by low-income households, and the appropriate databases to use to meet the particular needs of specific research projects. The most important of these recommendations are summarized in the final section of this chapter.

A. THE 1987-88 NFCS

Low rates of response to the 1987-88 NFCS by the households and individuals that comprised the survey's initial basic and low-income samples have resulted in small sample sizes and a high likelihood of bias in estimates based upon the survey databases.

1. Background

Small Sample Sizes. As planned, the basic and low-income samples of the 1987-88 NFCS were to have been roughly one-half to two-thirds the sizes of the actual basic and low-income samples of the 1977-78 NFCS. Because of low response rates, the actual samples of the 1987-88 NFCS are only 70 to 81 percent as large as planned. Therefore, the sizes of the actual basic and low-income samples of the 1987-88 NFCS are very much smaller than their counterparts in the 1977-78 NFCS.¹ In analyses of FNS programs, such as the FSP, that are targeted at broad segments of the population, the small sizes of the 1987-88 NFCS samples might result in only a modest loss of precision in

¹The actual sample sizes for the 1977-78 NFCS are as follows: *basic sample*--14,930 households completed the household interview, and 30,770 individuals provided at least one day of food intake data; *low-income sample*--4,623 households completed the household interview, and 12,847 individuals provided at least one day of food intake data. The actual sample sizes for the 1987-88 NFCS are as follows: *basic sample*--4,495 households completed the household interview, and 10,342 individuals provided at least one day of food intake data; *low-income sample*--2,509 households completed the household interview, and 6,443 individuals provided at least one day of food intake data.

estimates and would therefore be of little concern. In analyses of programs, such as WIC, that are targeted at narrow segments of the population, the small sample sizes could result in estimates that are not precise enough to detect true program impacts of moderate size. They could even preclude the possibility of analyzing the programs using the 1987-88 NFCS data.

Nonresponse Bias. If nonrespondents to the 1987-88 NFCS differ from respondents in their food consumption, then the high rates of nonresponse to the survey signal the existence of significant bias in estimates of food consumption behavior that are based on the survey data. We cannot fully assess the extent of nonresponse bias because no data are available on the food consumption behavior of nonrespondents. However, the LSRO expert panel found that respondents to the 1987-88 NFCS (basic sample) differ significantly from respondents to the March 1987 CPS in six demographic characteristics that are likely to influence food consumption. Even after the basic sample was weighted to correct for the over- or under-representation of selected demographic groups, the expert panel found the intake of food energy as reported in the 1987-88 NFCS to be significantly lower than that reported in the 1985 and 1986 CSFII.

2. Recommendations

Because of the high potential for nonresponse bias in estimates based on even the weighted 1987-88 NFCS data, we recommend that FNS avoid using these data for any analyses that will inform policymakers or update parameters in FNS' programs. Any such analyses would be intensely scrutinized by outside experts, and it would be difficult for FNS to defend the analyses.

The 1987-88 NFCS data can be used in analyses where the substantive findings are not intended to inform policymakers or update program parameters and, indeed, are secondary in importance to the development of the research methodology. The development of advanced statistical software to be used in future analyses of the impacts of multiple food assistance programs is an example of this type of appropriate application of the data.

B. ALTERNATIVE EXISTING OR PLANNED DATABASES

The 1987-88 NFCS is the only recent survey that has collected data on all three of the measures of food consumption that are commonly used as outcome measures in FNS-sponsored analyses of food consumption behavior. These measures are (1) expenditure on food by the household, (2) use of food at home by the household, and (3) food eaten ("food intake") by the individual. Many other recently completed or planned surveys have collected data on either one or two of these measures or on other measures of food consumption or related behavior. Because they provide only limited information on food consumption, none of these surveys is a good comprehensive substitute for the 1987-88 NFCS in the full range of its intended applications. However, in specific applications, one or more of these surveys used independently or in conjunction with each other may adequately substitute for the NFCS.

Our general recommendation to FNS regarding these alternative databases is that it plan to use them instead of the 1987-88 NFCS in many of the analyses of food consumption that it will sponsor over the next several years. Our specific recommendations regarding the use of each of these databases follow below.²

1. The 1989-91 CSFII

Background. The 1989-91 CSFII collects information on household expenditures on food, and more importantly, it is the only recent survey other than the 1987-88 NFCS to collect multiple days of data on food intake by individuals. The latter feature means that the CSFII is the only alternative

²We do not include SIPP and DHKS in our discussion of alternative databases that provide information on food consumption because, in fact, neither of these databases provides any information on food consumption. SIPP provides highly detailed information on program eligibility and participation and, hence, is an excellent source of data for modeling participation in USDA programs. From the main meal planner or preparer in each household that participated in the 1989-91 CSFII, the DHKS collects information on cooking practices, knowledge of the USDA/DHHS dietary guidelines, and knowledge of relationships between diet and health. The DHKS does not collect data on food consumption, but it can be linked to the CSFII, which does collect data on food consumption.

database that can be used to estimate the adjustment factors that the National Research Council (1986) argues are required for unbiased estimates of the proportion of a population group that is at nutritional risk.

Two important limitations of the 1989-91 CSFII are the small sizes of its basic and low-income samples and the absence of data on household food use.³ A third important limitation of the CSFII is that the income-eligibility limit for the WIC Program is 185 percent of poverty, whereas the survey's low-income samples include only households and individuals with incomes no greater than 130 percent of poverty. This incongruity between the CSFII and the analytic needs of the WIC program significantly reduces the usefulness of the survey to FNS.

Recommendations. To reduce the analytic limitations imposed by small sample sizes for any single year of the CSFII, we recommend combining the CSFII data across years for the basic sample and doing the same for the low-income sample. The combined samples will still be smaller than their counterparts in the 1987-88 NFCS, but they will be large enough to support analyses that are not focused on narrowly defined population groups. It will also be possible to use the combined data to estimate NRC adjustment factors that can then be used in analyses of other databases, such as NHANES III, that provide only a single day of food intake data but do so for many more individuals than does the CSFII.

To align the low-income samples of the CSFII with the analytic requirement of the WIC program, we recommend expanding the income cutoff for the low-income samples of the pending 1993-96 CSFII to 185 percent of poverty. FNS should also consider combining data from the basic and low-income samples of the 1989-91 CSFII by year for one or more years of the survey.

³Preliminary estimates of the sample sizes for the 1989 CSFII are as follows: *basic sample*--1,490 households completed the household interview, and 3,502 individuals provided at least one day of food intake data; *Low-income sample*--725 households completed the household interview, and 1,648 individuals provided at least one day of food intake data. The corresponding samples for the 1990 and 1991 CSFIIs are expected to be roughly the same size as the 1989 samples.

2. NHANES III

Background. From 1988 through 1994, NHANES III will collect one day of food intake data on 35,000 (projected) individuals ages 2 months and older. It will also obtain scientific measures of the health status of these individuals as well as information on economic and demographic characteristics and participation in assistance programs including the FSP, WIC, the SBP, and the NSLP. The survey does not have a separate low-income sample, but it does oversample several at-risk groups. The absence of multiple days of food intake data in NHANES III means that data from this survey alone cannot be used to estimate the distribution of usual daily intake of a given nutrient across the members of a population group or the proportion of that group that is at nutritional risk because of low intake of the nutrient.

Recommendations. We recommend that FNS use the NHANES III database in analyses of food and nutrient intake for which the 1989-91 CSFII sample sizes are inadequate or for which the CSFII low-income sample is inappropriate because its income ceiling is too low. One or both of these conditions are likely to arise in analyses of the WIC program and the school nutrition programs, especially in analyses of the effects of these programs on subgroups of their target populations, such as pregnant or breastfeeding women.

When estimates are required of the distribution of usual daily intake of a nutrient across the members of a population group, or if the proportion of that group that is at nutritional risk, the NHANES III data should be supplemented with estimates of adjustment factors, as recommended by the National Research Council (1986), obtained on the basis of the three days of intake data that are available in the 1989-91 CSFII. With these adjustment factors, the NHANES III data can be used to obtain the desired estimates.

3. The CES

Background. For between 5,000 and 6,000 consumer units each year, the CES Interview Survey obtains up to five observations on usual aggregate expenditures on food purchased for home

consumption and on food purchased and consumed away from home. For the same consumer units, the CES Diary Survey obtains data on expenditures on foods purchased during two consecutive one-week periods. The Diary Survey distinguishes between food purchased for consumption at home and food purchased for consumption away from home. Both surveys in the CES collect information on participation in the FSP and the school nutrition programs, but not in WIC. Neither of the surveys currently provides information on the quantity of food purchased or on the nutrients provided by that food.

Recommendations. We recommend that FNS consider using the CES Diary Survey as a source of data for analyses in which a measure of food expenditure is an appropriate outcome variable and in which a large sample of low-income consumer units is not required. Many analyses of the FSP would satisfy these conditions if FNS were willing to use a measure of food purchases rather than a measure of food use as the dependent variable. At least three studies of the effects of food stamp benefits on food consumption have been based on food purchase data from the CES Diary Survey (Salathe, 1980; Chavas and Yeung, 1982; West, 1984).

4. The PSID

Background. The PSID is an annual longitudinal survey, which in recent years has been obtaining data on about 7,100 families. Low-income families are overrepresented in the PSID sample. The survey currently collects information on family annual income, participation in the FSP (but not in any other USDA programs), and aggregate expenditure on food for home consumption and on food bought and eaten away from home.

Recommendations. Our recommendation to FNS regarding the use of the PSID in analyses of food consumption is essentially the same as our recommendation regarding the use of the CES. Senauer and Young (1986) have demonstrated that the PSID can support useful analyses of the effects of food stamp benefits on food consumption. We recommend that FNS consider basing such

analyses on the PSID when food consumption as measured by food expenditure is acceptable as an outcome variable.

5. The Cashout Demonstration Evaluation Databases

Background. In 1990, a uniform survey instrument was used to collect detailed data on household income, demographics, participation in USDA programs, and the use of food at home by roughly equal numbers of recipients of food stamp coupons and food stamp checks in San Diego County, Alabama, and Washington State. The three surveys together provide food use data for 4,553 households.

Recommendations. The greatest value to FNS of the Cashout databases will be realized in their capacity to support continued analysis and modeling of the effects of food stamp checks and coupons on the money value of food used by recipient households and on nutrient availability. To facilitate these analyses, we recommend that FNS combine the data from the three surveys into a single database. A combined Cashout database would, for example, make it easier for analysts to test for differences in the effects of cashout across the three demonstration sites.

The utility of these databases in analyses not related to cashout is limited because half of the households in the Cashout samples did not receive food stamp coupons, the samples include no eligible nonparticipants in the FSP, and the samples are not representative of the national food stamp caseload. Nevertheless, these databases should be useful in certain basic research applications, not related to cashout, in which the findings would not be generalized to the national food stamp caseload. The literature on the effects of food stamps on food consumption includes at least four studies based on data sets that were restricted to a single county (Neenan and Davis, 1977) or a single state (West and Price, 1976; West, Price, and Price, 1978; Price, West, Scheier, and Price, 1978).

6. Grocery Store Scan Data

Background. Grocery store checkout scanner systems have the potential to produce large amounts of highly disaggregated data on food purchases by food stamp recipients. Furthermore, they are likely to be able to do so in a timely manner while imposing virtually no response burden on the food stamp recipients. Very little of this potential has been realized to date, but advances in the storage and analysis of scanner data are rapidly being made by the retail food industry, commercial vendors of scanner data, and academic researchers.

Current deficiencies in scan data that severely detract from their utility to FNS are (1) the absence of UPC codes for variable-weight food items, which account for about 40 percent of retail food sales; (2) the absence of scanner systems in approximately 10 percent of supermarkets and in much larger percentages of smaller grocery stores; (3) the inaccessibility of information on sample design, on the treatment of missing data, and on sample weighting schemes for commercially available databases that are derived from scanner systems; (4) the apparently universal absence of scanner system files that contain fully disaggregated data on each transaction;⁴ (5) the only recent development of scanner system files that contain data aggregated over time on the specific food items purchased by individual "frequent shoppers," with linkages to files that provide information on the economic and demographic characteristics of these shoppers and their families.

Recommendations. We believe that checkout scanner systems will be a useful source of data on food consumption for FNS in the not too distant future. To avoid delays in taking advantage of these data once they do become available in a useful form, we recommend that FNS take steps now to become more familiar with checkout scanner systems, their associated databases, and emerging developments in the analysis of scan data. Three such steps would be for FNS to (1) keep up to date on scan data research developments by attending (and perhaps making presentations at) conference

⁴By "fully disaggregated data on each transaction," we mean data on each individual item purchased by each customer on each separate shopping trip.

sessions devoted to this topic and, more generally, by developing contacts with academics who are conducting research on scanner systems and scan data, (3) use FNS' small grants program as a vehicle for sponsoring research on scan data that more directly addresses FNS' needs than does most current academic research on this topic, and (3) acquire one of the commercially available scan databases and conduct some simple research on it in order to assess its potential and limitations.

7. The Database for the School Nutrition Dietary Assessment Study

Background. From a nationally representative sample of 3,500 students in grades 1 through 12, FNS' current School Nutrition Dietary Assessment (SNDA) Study is collecting data on eligibility for and participation in the SBP and NSLP, on dietary intake over a 24-hour period, and on other characteristics of the students, their schools, and their families.

Recommendations. Data generated by evaluations of programs or demonstrations have often proven useful to FNS in subsequent analyses that were outside the scope of the original evaluations. For example, FNS' current models of participation in school nutrition programs are based on data collected as part of the second National Evaluation of School Nutrition Programs (The Urban Institute, 1988). Data from the SNDA Study are also likely to have useful applications beyond their current scope. We recommend that FNS remain alert to the potential of these data to support future analyses of dietary intake and program participation by school-age children.

C. A NEW NATIONWIDE SURVEY OF FOOD USE BY LOW-INCOME HOUSEHOLDS

Much of the research on food consumption for which FNS has awarded contracts, but which is on hold pending a decision regarding the use of data from the 1987-88 NFCS, was intended to be conducted on household food use data (as opposed to data on either household food expenditures or individual food intake). This planned research includes such tasks as estimating the impacts of the FSP on nutrient availability and other measures of household food use, estimating economies of scale in household food consumption, and describing the food choices of low-income households.

Furthermore, HNIS had planned to use data on household food use in conjunction with data on individual food intake to revise the Thrifty Food Plan.

These are important research tasks; the findings would most likely lead to revisions in several critical parameters in the FSP. In Section A, we recommended that FNS avoid using the 1987-88 NFCS, which is the only recent source of nationally representative data on household food use, in such applications as these. That recommendation, along with the need to move forward with these research tasks, leads us now to recommend that FNS sponsor a new nationwide survey of food use by low-income households.

We recommend that the new survey be fielded as soon as possible and that it meet the following requirements:

1. The sample for the survey should be representative of households in the contiguous 48 states and the District of Columbia with monthly incomes no greater than 130 percent of poverty.⁵
2. The survey should complete interviews with approximately 5,000 households.⁶
3. The survey should collect data on household demographics, income, program participation and benefits, use of food at home, proportion of meals eaten at home, and expenditure on food used away from home.
4. To minimize respondent burden and thus maximize response rates, the survey should *not* collect data on individual food intake. Given the extreme importance

⁵We earlier recommended that the income cutoff for the 1993-96 CSFII be set at 185 percent of poverty rather than 130 percent. The reason for that recommendation is that the CSFII is, with the exception of the income screen, well-designed for research on the effects of the WIC and school nutrition programs on food and nutrient intake by individuals. However, the income-eligibility limit for these programs is 185 percent of poverty, so it would be highly desirable for the CSFII income cutoff to be set at that same level. In contrast, the new survey that we are proposing here would be a survey of food use by households. Its purpose would be to provide data that are required to revise the TFP and to conduct research on food use by food stamp recipients and income-eligible nonrecipients. Because the principal application of data collected by this survey would be to support research on the FSP, we recommend restricting the sample for the survey to households that satisfy the FSP gross income screen, which is set at 130 percent of poverty.

⁶Estimating the overall impact of the FSP on household food use would require data on only about 3,000 households, whereas revising the TFP would require data on about 5,000 households. The latter task would require a larger sample because it is necessary to analyze food use separately for each of the 31 TFP food groups.

of obtaining a high response rate on such a survey, we further recommend that the survey include only those questions that are essential to meeting FNS' need to conduct food use analyses in the context of the FSP and HNIS' need to revise the TFP.⁷

There exists a precedent for conducting a major survey of household food use that is independent of the NFCS. Following the implementation of major changes to the FSP in 1979, the SFC-LI was fielded in 1979-80 to obtain data on food consumption under the revised program regulations. This survey collected data on food use by about 3,000 households and data on one day of food intake by about 9,000 individuals in those households. We recommend against the collection of any intake data in the survey that we have proposed.

The respondent burden of the proposed survey would be similar to that of the surveys that were conducted in conjunction with the evaluations of the Food Stamp Cashout Demonstrations in Washington State, Alabama, and San Diego County. The fact that these surveys achieved response rates of 75 to 78 percent demonstrates that it is possible to attain high response rates on a survey with this level of respondent burden. Nevertheless, generating a high response rate to the proposed survey would be more challenging than on the Cashout surveys because the proposed survey would be fielded nationwide, whereas the Cashout surveys were fielded in a single county or in selected counties within a single state.

Estimating the cost of conducting the proposed survey is beyond the scope of the current project. However, we wish to point out that in addition to the obvious costs, such as those that would be incurred by the survey contractor and the government's internal costs of awarding and monitoring the contract, there would be less obvious, but nevertheless substantial, costs. For example, there would be the cost of updating HNIS' technical support system for conducting such a survey and for

⁷If essentially the same methodology that was used to develop the existing TFP were used to revise it on the basis of more recent data, then data on food intake by individuals would be required in addition to data on food use by individuals. However, it would not be necessary for both types of data to come from the same survey. We recommend that any intake data that may be needed to update the TFP be obtained either from the combined low-income samples of the 1989-91 CSFII or from NHANES III.

processing the resultant data. Because of the recent rapid pace of change in the retail food market, it should not be assumed that minor adjustments to the technical support system used by HNIS on the 1987-88 NFCS would be adequate.

D. RECOMMENDATIONS ON WAYS TO MEET THE DATA NEEDS OF SPECIFIC FNS RESEARCH PROJECTS

In Chapter II, we described five general research areas within FNS' overall research agenda that require survey data on food consumption:

1. Evaluating the impact of FNS' programs on food consumption
2. Designing FNS' food assistance programs
3. Modeling participation in FNS' food assistance programs
4. Describing the food consumption of the low-income population
5. Developing new research methodologies

Within each general area, we identified several specific research projects in which FNS has expressed interest. As conceptualized or designed, most of these projects would be based on data from the 1987-88 NFCS. In Section A, we recommended against using this database on all projects except those in which the development of research methodologies is of primary importance. Table VI.1 presents our recommendations on ways to meet the data needs of the specific FNS research projects we have identified.

E. SUMMARY OF RECOMMENDATIONS

In this chapter, we have made many recommendations regarding the use of the 1987-88 NFCS, the use of alternative databases, and ways to meet the food consumption data needs of specific FNS research projects. In order of importance, we summarize the most significant of those recommendations below.

1. We do not recommend using the 1987-88 NFCS on any research project that will inform policymakers or update parameters in FNS' programs.
2. We recommend conducting a new nationwide survey of food use by low-income households. There should be no individual intake component to this survey.⁸ The survey instrument should be designed with careful attention to the need to minimize respondent burden and hence to maximize response rates. The completed sample size should be approximately 5,000 households.⁹
3. We recommend setting the income ceiling for the low-income sample of the 1993-96 CSFII at 185 percent of poverty to make it consistent with the income-eligibility criteria for the WIC program and the school nutrition programs.
4. We recommend combining the samples of the 1989-91 CSFII across the three years of the survey and using the resultant databases on research projects that require data on food intake by individuals, especially those that require multiple days of intake data.
5. We recommend using NHANES III when it is necessary to analyze the food intake of individuals in small segments of the population.
6. We recommend using multiple databases on certain research projects. For example, we recommend using data from the 1989-91 CSFII and data from the recommended survey of household food use to (Recommendation 2, above) revise the Thrifty Food Plan.
7. We recommend using data from program and demonstration evaluations in analyses other than those required by the evaluations. For example, FNS' model of participation in the National School Lunch Program is based on data from the National Evaluation of School Nutrition Programs. Opportunities for analogous applications of the data from the Food Stamp Cashout Studies and the School Nutrition Dietary Assessment Study are likely to arise in the future.
8. We recommend that FNS enhance its knowledge of grocery store checkout scanner systems and their associated databases and that it monitor ongoing developments in the collection, storage, dissemination, and analysis of such data.

⁸The CSFII provides adequate data on individual intake for most FNS applications.

⁹A sample of 5,000 households would be adequate for revising the Thrifty Food Plan and would be more than adequate for analyses of the effects of the FSP on household food use.

TABLE VI.1

**EXAMPLES OF FNS' RESEARCH PROJECTS AND RECOMMENDATIONS
ON WAYS TO MEET THEIR DATA NEEDS**

Research Project	Recommendations
A. Evaluating the Impact of FNS' Programs on Food Consumption	
1. Estimating the Impact of the FSP on Household Food Use	<ul style="list-style-type: none"> • Conduct a new household food use survey of about 3,000 low-income households.
2. Estimating the Impact of WIC on the Diets of Women, Infants, and Children	<ul style="list-style-type: none"> • Use NHANES III if estimates of the impacts of WIC on narrowly defined segments of the WIC-eligible population are required. • If objective is to estimate proportions of the WIC-eligibles at nutritional risk, then use combined data from basic samples of 1989-91 CSFII to estimate NRC adjustment factors; small sample size will be a problem for some subgroups. Use adjustment factors in conjunction with data from combined CSFIs or NHANES III.
3. Evaluating the Effectiveness of Nutrition Education Programs	<ul style="list-style-type: none"> • If program was implemented or expanded nationwide during 1989-91, then link DHKS and CSFII for each year in period. • If program is small in scale, then conduct new survey of participants and nonparticipants in the vicinity of program site. Include questions about diet, participation in nutrition education programs, and diet-related knowledge and attitudes.
B. Program Design	
1. Revising the Thrifty Food Plan	<ul style="list-style-type: none"> • Conduct a new household food use survey of about 4,500 to 5,000 low-income households. • Obtain individual food intake data from combined low-income samples of 1989-91 CSFII or from NHANES III.
2. Estimating Economies-of-Scale Adjustments in FSP benefits	<ul style="list-style-type: none"> • Conduct a new household food use survey of about 4,000 low-income households.
3. Changing (1) the FSP Benefit Issuance System, (2) Criteria for Authorizing Retail Stores, and (3) Coupon Denominations	<ul style="list-style-type: none"> • Investigate grocery store scan data. This may entail contracting with (1) supermarkets to store data on individual transactions, and/or (2) commercial firms that collect household food purchase data to also obtain information on FSP participation.
C. Modeling Program Participation	
1. General Modeling of Participation in FNS' Programs	<ul style="list-style-type: none"> • Use SIPP data.
2. Modeling Participation in the NSLP and SBP	<ul style="list-style-type: none"> • Use data from the SNDA Study.
D. Descriptive Studies of Food Consumption	
1. Describing the Food Choices of Low-Income Households	<ul style="list-style-type: none"> • Conduct a new household food use survey of about 3,000 low-income households.
2. Describing the Food Intake of Low-Income Children	<ul style="list-style-type: none"> • Use a variety of databases, including the database from the SNDA Study, NHANES III, and the combined basic samples of 1989-91 CSFII.

TABLE VI.1 (continued)

Research Project	Recommendations
E. Methodological Studies	
1. Estimating the Distribution of Usual Dietary Intake with Only One Day of Dietary Intake	<ul style="list-style-type: none"> • Use combined basic or low-income samples of 1989-91 CSFII to estimate NRC adjustment factors; small sample size will be a problem for some subgroups. Use adjustment factors in conjunction with data from combined CSFIIs or NHANES III.
2. Developing Methodologies for Analyzing Self-Selection in Models of Multiple Program Participation	<ul style="list-style-type: none"> • Since the substantive results are secondary in importance to the model development, virtually any database would be acceptable so long as it provides (1) large numbers of observations on persons or households who are eligible for multiple assistance programs, and (2) a measure of food consumption or some other outcome measure of interest to FNS.
3. Validating Data from the Food Stamp Cashout Demonstration Evaluations	<ul style="list-style-type: none"> • Use SIPP to make comparisons between the demographic and socioeconomic characteristics of FSP households in the Cashout sites and FSP households nationwide. • Use data from a new household food use survey of low-income households to make comparisons between the food use of FSP households in the Cashout sites (use data for coupon recipients only) and FSP households nationwide.

REFERENCES

- Allin, Susan, and Harold Beebout. "Determinants of Participation in the Food Stamp Program: A Review of the Literature." Washington, DC: Mathematica Policy Research, Inc., November 1989.
- Battese, G.E., S.M. Nusser, and W.A. Fuller. "Estimation of the Distribution of Usual Intakes for Selected Dietary Components Using Data From the 1985-86 Continuing Survey of Food Intakes by Individuals." Ames, IA: Center for Agricultural and Rural Development and Department of Statistics, Iowa State University, 1988.
- Beebout, Harold, Edward Cavin, Barbara Devaney, Thomas Fraker, Sharon Long, and Peter Mossel. "Evaluation of the Nutrition Assistance Program in Puerto Rico--Volume II: Effects on Food Expenditures and Diet Quality." Washington, DC: Mathematica Policy Research, Inc., 1985.
- Blanchard, Lois, J.S. Butler, Pat Doyle, Russell Jackson, James C. Ohls, and Barbara M. Posner. "Food Stamp SSI/Elderly Cashout Demonstration Evaluation." Princeton, NJ: Mathematica Policy Research, Inc., 1982.
- Burghardt, John, Merrile Sing, Barbara Devaney, Wendy Foran, and Charlene Weiss. "School Nutrition Dietary Assessment: Study Design and Analysis Plan." Princeton, NJ: Mathematica Policy Research, Inc., October 1990.
- Chavas, Jean-Paul, and M.L. Yeng. "Effects of the Food Stamp Program on Food Consumption in the Southern United States." *Southern Journal of Agricultural Economics*, vol. 14, 1982, pp. 131-139.
- Czajka, J.L. "Determinants of Participation in the Food Stamp Program: Spring 1979." Washington, DC: Mathematica Policy Research, Inc., 1981.
- Devaney, Barbara, Pamela Haines, and Robert Moffitt. "Assessing the Dietary Effects of the Food Stamp Program." Princeton, NJ: Mathematica Policy Research, Inc., February 1989.
- Devaney, Barbara, and Thomas Fraker. "The Effect of Food Stamps on Food Expenditures: An Assessment of Findings from the Nationwide Food Consumption Survey." *American Journal of Agricultural Economics*, vol. 71, February 1989, pp. 99-104.
- "44th Annual Consumer Expenditures Study." *Supermarket Business*, vol. 46, no. 9, September 1991, pp. 47-68.
- Fraker, Thomas M., Sharon K. Long, and Charles E. Post. "Analyses of the 1985 Continuing Survey of Food Intakes by Individuals, Volume I." Washington, DC: Mathematica Policy Research, Inc. September 1990.
- Fraker, Thomas M., Alberto P. Martini, James C. Ohls, Michael Ponza, and Elizabeth A. Quinn. "The Evaluation of the Alabama Food Stamp Cash-Out Demonstration, Volume I (Draft)." Washington, DC: Mathematica Policy Research, Inc., February 1992a.

- _____. "The Evaluation of the Alabama Food Stamp Cash-Out Demonstration, Volume II (Draft)." Washington, DC: Mathematica Policy Research, Inc., February 1992b.
- Fraker, Thomas M., and R. Moffitt. "The Effect of Food Stamps and Labor Supply: A Bivariate Selection Model." *Journal of Public Economics*, vol. 35, 1988, pp. 25-56.
- Fraker, Thomas M., and Charlie Post. "Documentation for a SAS Analysis File Containing Merged Wave 1 Data for the Core and Low-Income Samples of the 1985 CSFII." Washington, DC: Mathematica Policy Research, Inc., February 1987.
- _____. "Documentation for a SAS Analysis File Containing Merged 4-Day Data for the Core and Low-Income Samples of the 1985 CSFII." Washington, DC: Mathematica Policy Research, Inc., March 1989.
- General Accounting Office. "Nutrition Monitoring: Mismanagement of Nutrition Survey Has Resulted In Questionable Data." Washington DC, July 1991.
- Heckman, James J. "Dummy Endogenous Variables in a Simultaneous Equations System." *Econometrica*, vol. 46, July 1978, pp. 931-961.
- _____. "Sample Selection Bias as a Specification Error." *Econometrica*, vol. 47, January 1979, pp. 153-161.
- Heckman, James J., and Richard Robb. "Alternative Methods for Evaluating the Impact of Interventions," in *Longitudinal Analysis of Labor Market Data*, ed. by J. Heckman and B. Singer. Cambridge, England: Cambridge University Press, 1985.
- Hepburn, Frank N. "The USDA National Nutrient Data Bank." *The American Journal of Clinical Nutrition*, vol. 35, 1982, pp. 1297-1301.
- Hubbard, Michael, Emelita deLeon-Wong, Daniel Hungerford, Judith Lessler, and Rachel Caspar. "Sampling Designs and Population Descriptors of Nationwide Food Consumption Surveys and National Health and Nutrition Examination Surveys." Research Triangle Park, NC: Research Triangle Institute, May 1991.
- Human Nutrition Information Service. "The Thrifty Food Plan, 1983." Hyattsville, MD, August 1983.
- Kerr, Richard, and Betty Peterkin. "The Effect of Household Size on the Cost of Diets that are Nutritionally Equivalent." Unpublished Report. Hyattsville, MD: Consumer and Food Economics Institute, U.S. Department of Agriculture, September 1975.
- Kisker, Ellen Eliason, and Barbara Devaney. "The Food Choices of Low-Income Households." Princeton, NJ: Mathematica Policy Research, Inc., January 1988.
- Life Sciences Research Office. "Impact of Nonresponse on Dietary Data From the 1987-88 Nationwide Food Consumption Survey." Bethesda, MD: Federation of American Societies for Experimental Biology, April 1991.
- Long, Sharon K. "Do the School Nutrition Programs Supplement Household Food Expenditures?" *Journal of Human Resources*, vol. 26, no. 4, fall 1991, pp. 654-678.

- Loughin, Marie M., and Wayne A. Fuller. "Construction of Regression Weights For the Nationwide Food Consumption Survey Data." Ames, IA: Department of Statistics, Iowa State University, December 1990.
- Martini, Alberto. "Participation in the Food Stamp Program: A Multivariate Analysis." Washington, DC: Mathematica Policy Research, Inc., March 1992.
- McFadden, Daniel. "A Method of Simulated Moments for Estimation of Multinomial Probits Without Numerical Integration." Mimeograph. Boston: Massachusetts Institute of Technology, 1986.
- National Research Council, Committee on Dietary Allowances. *Recommended Dietary Allowances*. Washington, DC: National Academy Press, 1980.
- National Research Council, Subcommittee on Criteria for Dietary Evaluation. *Nutrient Adequacy: Assessment Using Food Consumption Surveys*. Washington, DC: National Academy Press, 1986.
- National Research Council, Committee on Dietary Allowances. *Recommended Dietary Allowances*. Washington, DC: National Academy Press, 1989.
- Neenan, Pamela H., and Carlton G. Davis. "Impact of the Food Stamp Program on Low Income Household Food Consumption in Rural Florida." *Southern Journal of Agricultural Economics*, vol. 9, 1977, pp. 89-97.
- Nelson, Lyle, Harold Beebout, and Felicity Skidmore. "Economies of Scale in the Food Stamp Program." Alexandria, VA: Food and Nutrition Service, U.S. Department of Agriculture, February 1985.
- Pakes, Ariel, and David Pollard. "The Asymptotics of Simulation Estimators." Working paper no. 8628. Madison, WI: Social Systems Research Institute, University of Wisconsin, September 1986.
- Pao, Eleanor M., Katherine E. Sykes, and Yasmin S. Cypel. "USDA Methodological Research for Large-Scale Dietary Intake Surveys, 1975-88." Hyattsville, MD: Human Nutrition Information Service, U.S. Department of Agriculture, December 1989.
- Price, David W., Donald A. West, Genevieve E. Schier, and Dorothy Z. Price. "Food Delivery Programs and Other Factors Affecting Nutrient Intake of Children." *American Journal of Agricultural Economics*, vol. 60, 1978, pp. 609-618.
- Ritenbaugh, Cheryl, et al. "Methodological Issues in Food Consumption Surveys." Final report under USDA cooperative agreement #58-3198-6-62, July 1988.
- Salathe, Larry E. "The Food Stamp Program and Low-Income Households' Food Purchases." *Agricultural Economics Research*, vol. 32, no. 4, 1980, pp. 33-41.
- Scearce, W. Keith, and Robert Jensen. "Food Stamp Program Effects on Availability of Food Nutrients for Low Income Families in the Southern Region of the United States." *Southern Journal of Agricultural Economics*, vol. 11, December 1979, pp. 113-120.

- Schoen, E. "Once Again, Hunger Troubles America." *New York Times Magazine*, January 2, 1983.
- The President's Private Sector Survey on Cost Control (the Grace Commission), USDA Task Force. *The Food and Nutrition Services Thrifty Food Plan Economies of Scale Issue and Recommendation Summaries*, No. AG-9 and AG-10, vol. 1, September 14, 1982.
- U.S. Department of Agriculture and Department of Health and Human Services. *Dietary Guidelines for Americans*. Washington, DC: U.S. Government Printing Office, 1980.
- U.S. Department of Agriculture and Department of Health and Human Services. *Dietary Guidelines for Americans*. Washington, DC: U.S. Government Printing Office, 1990.
- Uniform Code Council, Inc. *UPC Implementation Guide: How to Develop and Maintain a Top Quality UPC*. Dayton, OH: Uniform Code Council, Inc., 1991.
- The Urban Institute. "Modeling Student Participation in School Nutrition Programs." Washington, DC, March 1988.
- West, Donald A. "Effects of the Food Stamp Program on Food Expenditures." Agricultural Research Center, Washington State University, Research Bulletin XB 0922, 1984.
- West, Donald A., and David W. Price. "The Effects of Income, Assets, Food Programs, and Household Size on Food Consumption." *American Journal of Agricultural Economics*, vol. 58, 1976, pp. 725-730.
- West, Donald A., David W. Price, and Dorothy Z. Price. "Impacts of the Food Stamp Program on Value of Food Consumed and Nutrient Intake among Washington Households with 8-12 Year Old Children." *Western Journal of Agricultural Economics*, vol. 3, 1978, pp. 131-144.

APPENDIX

ABSTRACTS OF FOOD CONSUMPTION DATABASES

APPENDIX CONTENTS

	Page
THE 1987-88 NATIONWIDE FOOD CONSUMPTION SURVEY (NFCS)	A-5
THE 1989-91 CONTINUING SURVEY OF FOOD INTAKES BY INDIVIDUALS (CSFII)	A-9
THE NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY III (NHANES III)	A-11
THE CONSUMER EXPENDITURE SURVEY (CES)	A-13
THE 1977-78 NATIONWIDE FOOD CONSUMPTION SURVEY (NFCS)	A-17
THE 1979-80 SURVEY OF FOOD CONSUMPTION IN LOW-INCOME HOUSEHOLDS (SFC-LI)	A-21
THE SURVEY OF INCOME AND PROGRAM PARTICIPATION (SIPP)	A-23
THE PANEL STUDY OF INCOME DYNAMICS (PSID)	A-25
THE DIET AND HEALTH KNOWLEDGE SURVEY (DHKS)	A-27
THE FOOD STAMP CASHOUT EVALUATION SURVEYS	A-29
THE SCHOOL NUTRITION DIETARY ASSESSMENT STUDY	A-31

THE 1987-88 NATIONWIDE FOOD CONSUMPTION SURVEY (NFCS)

Sponsoring Agency

USDA/HNIS

Target Population

Private households in the 48 contiguous states and all resident individuals within those households

1. Basic sample: households with all levels of income
2. Low-income sample: households with income at or below 130 percent of poverty

Data Collection Period

April 1987 to August 1988

Unit of Observation--Food Consumption Data

Household and individual

Food Consumption Measures

1. Household
 - Expenditure per week or month during the previous three months on food for home consumption
 - Expenditure during the seven previous days on food bought and eaten away from home
 - Food used from home supplies during the previous seven days--type, form (frozen, canned, etc.), quantity bought, price, quantity used, money value used, source (bought, home-produced, etc.)
2. Individual
 - Food eaten at home and away from home for three consecutive days

Method of Collection--Food Consumption Data

In-person interviews

1. Household
 - Aided seven-day recall
2. Individual
 - One-day recall (previous day) and two-day record (day of interview and next day)

Income Measures

Previous month's income, by source, for each household member aged 15 or older, and combined household income over the previous calendar year, not by source

Sample Size

1. Basic sample
 - 4,495 households completed the household interview
 - 10,342 individuals completed at least one day of intake data
 - 8,468 individuals completed all three days of intake data
2. Low-income sample
 - 2,509 households completed the household interview
 - 6,443 individuals completed at least one day of intake data (preliminary)
 - 5,309 individuals completed all three days of intake data (preliminary)

Response Rate

1. Basic sample
 - 38 percent of households contacted completed the household interview
 - 31 percent of individuals in the households contacted completed at least one day of intake data
 - 26 percent of individuals in the households contacted completed all three days of intake data

2. Low-income sample
 - Not yet available

Measures of Program Participation

1. FSP: whether any household member received food stamps in the previous 12 months, whether household currently receives food stamps, names of recipients, date and amount of most recent benefit
2. WIC: whether any household member received WIC benefits in the previous month, name of each recipient
3. NSLP and SBP: whether any child in the household receives school lunches/breakfasts during the school year, name of each recipient, number of times a week each recipient gets school lunches/breakfasts, meal-price status (free/paid) of each recipient, amount spent on each lunch/breakfast, meal prices as reported by School Food Authority--via a link to the School Nutrition Programs Update database

Availability of Data

1. Basic sample
 - Currently available
2. Low-income sample
 - Household data will be available in the spring of 1992

THE 1989-91 CONTINUING SURVEY OF FOOD INTAKES BY INDIVIDUALS (CSFII)

Sponsoring Agency

USDA/HNIS

Target Population

Private households in the 48 contiguous states and all resident individuals within those households

1. Basic sample: households with all levels of income
2. Low-income sample: households with income at or below 130 percent of poverty

Data Collection Period

April to March for each year between 1989-1991

Unit of Observation--Food Consumption Data

Household and individual

Food Consumption Measures

1. Household
 - Expenditure per week or month during the previous three months on food for home consumption
 - Expenditure per week or month during the previous three months on food bought and eaten away from home
2. Individual
 - Food eaten at home and away from home for three consecutive days

Method of Collection--Food Consumption Data

In-person interview, one-day recall (previous day) and two-day record (day of interview and next day)

Income Measures

Combined household income during the previous month, by source, and during the previous calendar year, not by source

Sample Size

1. Basic sample

- 1,490 households completed the household interview in 1989 (preliminary)
- 3,502 individuals completed at least one day of intake data (preliminary)
- 2,925 individuals completed all three days of intake data (preliminary)

2. Low-income sample

- 725 households completed the household interview in 1989 (preliminary)
- 1,648 individuals completed at least one day of intake data (preliminary)
- 1,372 individuals completed all three days of intake data (preliminary)

Response Rate

Not available

Measures of Program Participation

1. FSP: whether any household member received food stamps in the previous 12 months, whether household currently receives food stamps, names of recipients, date and amount of most recent benefit
2. WIC: whether any household member currently receives WIC benefits, name of each recipient, number of months receiving benefits
3. NSLP and SBP: whether any child in the household receives school lunches/breakfasts during the school year, name of each recipient, number of times a week each recipient gets school lunches/breakfasts, meal-price status (free/reduced-price/full-price) of each recipient

Availability of Data

1989 data will be available in the spring of 1992

1990 data will be available in the fall of 1992

1991 data will be available in the summer of 1993

THE NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY III (NHANES III)

Sponsoring Agency

DHHS\CDC\NCHS

Target Population

Civilian, noninstitutionalized persons, aged 2 months and older, in 50 states and the District of Columbia

Data Collection Period

1. Phase 1: 1988-91
2. Phase 2: 1992-94

Unit of Observation--Food Consumption Data

Individual

Food Consumption Measures

1. Food eaten at home and away from home during a 24-hour period
2. Frequency of food items eaten in the previous month

Method of Collection--Food Consumption Data

In-person interview, one-day recall

Income Measures

Combined family income during the previous month, by source, and during the previous 12 months, not by source

Sample Size

- 30,000 persons examined (projected)
- 35,000 persons interviewed (projected)

Response Rate

As of April 1991, 77 percent of the individuals who had been contacted had completed both the interview and the examination components of the survey.

Measures of Program Participation

1. FSP: whether any household member received food stamps in the previous 12 months and whether any household member is currently receiving food stamps
2. WIC: household--whether any household member received benefits in previous month; women and children--whether received benefits in the previous 12 months, whether now receiving benefits, length of time received benefits
3. NSLP and SBP: number of times a week a child usually receives school lunches/breakfasts

Availability of Data

Data from phase 1 of the survey will be available in 1992

THE CONSUMER EXPENDITURE SURVEY (CES)

Sponsoring Agency

U.S. Department of Labor/BLS

Target Population

Civilian noninstitutionalized persons and persons living in group quarters who reside in the 50 states and the District of Columbia

Data Collection Period

Continuous

1. Interview

- Respondents interviewed every quarter for five quarters

2. Diary

- Respondents keep diary for two consecutive one-week periods

Unit of Observation--Food Consumption Data

Consumer unit--all members of a household who are related by blood, marriage, adoption, or other legal arrangements; persons living alone or sharing a household with others; two or more persons who live together and who pool their income to make joint expenditure decisions

Food Consumption Measures

1. Interview

- Usual expenditure per week during the previous three months on food for home consumption
- Usual expenditure per month during the previous three months on food bought and eaten away from home

2. Diary

- Usual expenditure per week during the previous three months on food for home consumption
- Expenditure over two consecutive one-week periods on food bought and eaten away from home--meal occasion, total cost
- Expenditure over two consecutive one-week periods on food for home consumption--type, form (frozen, canned, etc.), total cost. 1990 was the last year that food quantity data were collected.

Method of Collection--Food Consumption Data

In-person interview

1. Interview

- Aided three-month recall

2. Diary

- Two one-week records, reviewed by interviewer

Income Measures

1. Interview

- Income over the previous 12 months for each person aged 14 and older in the consumer unit, by source

2. Diary

- Income over the previous 12 months for each person aged 14 and older in the consumer unit, by source

Sample Size

1. Interview

- 20,517 observations (approximately four observations per consumer unit)

2. Diary

- 11,735 observations (approximately two observations per consumer unit)

Response Rate

1. Interview

- 85.7 percent of eligible consumer units were interviewed in 1990

2. Diary

- 86.6 percent of eligible consumer units were interviewed in 1990

Measures of Program Participation

1. Interview

- FSP: number of months household received food stamps over the previous three months, benefit amount
- WIC: none
- NSLP and SBP: whether any child in the household purchased school meals during previous three months, name of child, child's usual weekly expense on school meals

2. Diary

- FSP: number of months household received food stamps over the previous 12 months, whether household received food stamps in the previous month, benefit amount in previous month
- WIC: none
- NSLP and SBP: whether any child in the household purchased school meals during previous 30 days, name of child, child's usual weekly expense on school meals

Availability of Data

1. 1990 data are currently available
2. 1991 data will be available in the fall of 1992

THE 1977-78 NATIONWIDE FOOD CONSUMPTION SURVEY (NFCS)

Sponsoring Agency

USDA/HNIS

Target Population

Private households in the 48 contiguous states that are either receiving food stamps or other welfare assistance or have income and assets low enough to meet the FSP eligibility requirements and all resident individuals in those households

Data Collection Period

April 1977 to March 1978

Unit of Observation--Food Consumption Data

Household and individual

Food Consumption Measures

1. Household
 - Expenditure per week or month during the previous three months on food for home consumption
 - Expenditure during the previous seven days on food bought and eaten away from home
 - Food used from home supplies during the previous seven days--type, form (frozen, canned, etc.), quantity bought, price, quantity used, money value used, source (bought, home-produced, etc.)
2. Individual
 - Food eaten at home and away from home during a 24-hour period

Method of Collection--Food Consumption Data

In-person interview

1. Household
 - Aided seven-day recall
2. Individual
 - One-day recall (previous day) and two-day record (day of interview and the next day)

Income Measures

Previous month's income, by source, for each household member aged 14 or older, and combined household income over the previous calendar year, not by source

Sample Size

1. Basic sample
 - 14,930 households completed the household interview
 - 30,770 individuals completed at least one day of intake data
2. Low-income sample
 - 4,623 households completed the household interview
 - 12,847 individuals completed at least one day of intake data

Response Rate

1. Basic sample
 - 61 percent of households contacted completed the household interview
 - 57 percent of individuals in households contacted completed at least one day of intake data
2. Low-income sample
 - Not available

Measures of Program Participation

1. FSP: previous 12 months--whether any household member received food stamps, number of months received food stamps; previous month--whether household received food stamps, number of household members authorized to receive food stamps, benefit amount; current month--whether household currently receives food stamps
2. WIC: whether any household member received WIC benefits in the previous month; number of women, infants, and children who received benefits
3. NSLP and SBP: whether any child in the household receives school lunches/breakfast during the school year, name of each recipient, number of times a week each recipient gets school lunches/breakfasts, amount spent on each lunch/breakfast

Availability of Data

Currently available

**THE 1979-80 SURVEY OF FOOD CONSUMPTION IN LOW-INCOME HOUSEHOLDS
(SFC-LI)**

Sponsoring Agency

USDA/HNIS

Target Population

Private households in the 48 contiguous states that are eligible to participate in the FSP and all resident individuals in those households

Data Collection Period

November 1979 to March 1980

Unit of Observation--Food Consumption Data

Household and individual

Food Consumption Measures

1. Household

- Expenditure per week or month during the previous three months on food for home consumption
- Expenditure during the previous seven days on food bought and eaten away from home
- Food used from home supplies during the previous seven days--type, form (frozen, canned, etc.), quantity bought, price, quantity used, money value used, source (bought, home-produced, etc.)

2. Individual

- Food eaten at home and away from home during a 24-hour period

Method of Collection--Food Consumption Data

In-person interview

1. Household
 - Aided seven-day recall
2. Individual
 - One-day recall

Income Measures

Previous month's income, by source, for each household member aged 14 or older, and combined household income over the previous calendar year, not by source

Sample Size

1. 3,002 households
2. 9,123 individuals

Response Rate

Not available

Measures of Program Participation

1. FSP: previous 12 months--whether any household member received food stamps, number of months received food stamps; previous month--whether household received food stamps, number of household members authorized to receive food stamps, benefit amount; current month--whether household currently receives food stamps
2. WIC: whether any household member received WIC benefits in the previous month; number of women, infants, and children who received benefits
3. NSLP and SBP: whether any child in the household receives school lunches/breakfast during the school year, name of each recipient, number of times a week each recipient gets school lunches/breakfasts, amount spent on each lunch/breakfast

Availability of Data

Currently available

THE SURVEY OF INCOME AND PROGRAM PARTICIPATION (SIPP)

Sponsoring Agency

U.S. Department of Commerce/Bureau of the Census

Target Population

Civilian, noninstitutionalized population in 50 states and the District of Columbia

Data Collection Period

1. Continuous since 1983; respondents are typically interviewed eight times at four-month intervals although some observations in the 1990 panel were interviewed 12 times
2. A new sample ("panel") is fielded in February of each year; data were collected for most households from the 1990 panel from February 1990 until May 1992 and collected from February 1989 until May 1992 for a carryover sample (defined below).

Unit of Observation

Household and individual

Food Consumption Measures

None

Method of Data Collection

In-person interviews

Income Measures

Income in each of the preceding four months, by source, for each household member aged 15 or older

Sample Size

17,748 households from the national probability sample and 3,458 households from the "carryover" sample (introduced to achieve an oversampling of blacks, Hispanics, and single female-headed households) were interviewed in Wave 1 of the 1990 panel.

Response Rate

1. Over 90 percent in the first wave of the 1990 panels for households introduced in February 1990
2. Over 80 percent in other waves of the 1990 panel

Measures of Program Participation

1. FSP: name of each household member who was authorized to receive food stamps in any of the previous four months, name of each household member covered under each authorized recipient's benefit, value of each authorized recipient's benefit in each of the previous four months
2. WIC: name of each household member who was covered by WIC benefits in any of the previous four months, months in which each covered person received benefits
3. NSLP and SBP: number of children in the household who receive free or reduced-price school lunches/breakfasts in the current school year, "household" meal-price status (free/reduced-price/full-price) if any meals were received

Availability of Data

Data are now available for all waves of the 1984-1988 panels and for the first two waves of the 1990 panel. There are no current plans to release data from the 1989 panel, which was abbreviated in order to introduce a larger more efficient sample for the 1990 panel.

THE PANEL STUDY OF INCOME DYNAMICS (PSID)

Sponsoring Agency

DHSS

Target Population

Families in the 50 states and the District of Columbia, low-income families were oversampled

Data Collection Period

Annually since 1968. Food consumption data were collected in 1968 through 1987 and in 1990 to the present. Food consumption data were not collected in 1988 and 1989.

Unit of Observation--Food Consumption Data

Family

Food Consumption Measures

1. Expenditure per week or month during previous month on food for home consumption, excluding purchases made with food stamps
2. Expenditure per week or month during previous month on food bought and eaten away from home, excluding meals at work or school

Method of Collection--Food Consumption Data

1. 1968-1972: primarily in-person interviews
2. 1973-1987: primarily telephone interviews
3. 1988-1989: food consumption data were not collected
4. 1990-present: primarily telephone interviews

Income Measures

Combined family income over the previous calendar year, by source

Sample Size

7,114 families in 1988

Response Rate

1. 1968 family response rate:
 - 76 percent of families contacted responded to the survey
2. 1988 family response rates:
 - 97 percent of families contacted in 1988 responded to the survey in 1988
 - 37 percent of families contacted in 1968 responded to the survey in 1988

Measures of Program Participation

1. FSP
 - 1968-1987 and 1990-present: previous calendar year--whether any family member received food stamps, number of months of reciprocity, monthly benefit amount; previous month--whether any family member received food stamps, number of recipients, benefit amount
 - 1988-1989: previous calendar year--whether any family member received food stamps, number of months of reciprocity, monthly benefit amount; previous month--no participation data were collected
2. WIC: participation data are available for 1985 only
3. NSLP and SBP: none

Availability of Data

1968-1988 data are currently available

1989 data will be available in March 1992

1990 data will be available in late 1992

THE DIET AND HEALTH KNOWLEDGE SURVEY (DHKS)

Sponsoring Agencies

USDA/HNIS and Food Safety and Inspection Service, and DHHS/FDA

Target Population

Households that responded to the basic and low-income components of the CSFII

Data Collection Period

10 weeks after the CSFII interviews in 1989-91

Food Consumption Measures

1. None, but can be linked to CSFII
2. Collects information on knowledge and attitudes about dietary guidance, food preparation practices, use of nutrition information on labels, and food safety concerns

Method of Collection

Telephone interview (in-person, if respondent cannot be contacted by telephone)

Unit of Observation

Main meal planner/preparer in household

Income Measures

Available from CSFII

Sample Size

1. 1989 Basic sample
 - 1,280 households (preliminary)

2. 1989 Low-income sample

- 626 households (preliminary)

Response Rate

In 1989, 86 percent of the households that responded to the CSFII also responded to the DHKS

Measures of Program Participation

Available from CSFII

Availability of Data

1989 data will be available in the summer of 1992

THE FOOD STAMP CASHOUT EVALUATION SURVEYS

Sponsoring Agency

USDA/FNS

Target Population

Households receiving check or coupon food stamp benefits in three sites:

1. San Diego County, California
2. Twelve counties in Alabama
3. Five check counties and five coupon counties in Washington State

Data Collection Period

1. San Diego County: May through August, 1990
2. Alabama: August through November, 1990
3. Washington State: August through October, 1990

Unit of Observation--Food Consumption Data

Household

Food Consumption Measures

1. Expenditure during the previous month on food for home consumption
2. Expenditure during the previous seven days on food bought and eaten away from home
3. Food used from home food supplies during the previous seven days--type, form (frozen, canned, etc.), quantity bought, price, quantity used, money value used, source (bought, home-produced, etc.)

Method of Collection--Food Consumption Data

In-person interview, aided seven-day recall

Income Measures

Previous month's income, by source, for each household member aged 14 or older

Sample Size

1. San Diego County: food use data for 1,078 households
2. Alabama: food use data for 2,291 households
3. Washington State: food use data for 1,184 households

Response Rate

1. San Diego County: 78 percent
2. Alabama: 78 percent
3. Washington State: 75 percent

Measures of Program Participation

1. FSP: names of persons covered by household's food stamp benefit, benefit amount in the previous month and current month
2. WIC: whether any household member received WIC benefits in the previous month, name of each recipient, benefit amount in the previous month
3. NSLP and SBP: total number of school lunches/breakfasts received in the previous month by all children in the household, "household" meal-price status (free/reduced-price/full-price) if any school meals were received, total amount spent in the previous month on school lunches/breakfasts

Availability of Data

1. San Diego County and Alabama: fall of 1992
2. Washington State: fall of 1992

THE SCHOOL NUTRITION DIETARY ASSESSMENT STUDY

Sponsoring Agency

USDA/FNS

Target Population

1. 350 nationally representative schools in the 48 contiguous states and the District of Columbia (school-level data will be collected from 225 additional schools)
2. 3,500 children in grades 1 through 12 that attend the 350 schools

Data Collection Period

January 1992 through May 1992

Unit of Observation--Food Consumption Data

Individual

Food Consumption Measures

Food consumed at home and away from home during a 24-hour period

Method of Collection--Food Consumption Data

1. In-person interview at school with children in grades 3 through 12, in-person interview at home with children in grades 1 and 2 and their parents
2. One-day recall

Income Measures

Annual combined household income

Sample Size

3,500 children (target sample size)

Response Rate

Not available

Measures of Program Participation

1. FSP: whether the parent(s) of child currently receive(s) food stamps
2. WIC: none
3. NSLP and SBP: whether the parent applied for free or reduced-price meals, number of days per week child eats school lunch, meal-price status of each child

Availability of Data

Fall of 1992